UNIVERSITY OF TORONTO
3 1761 00465594 0

ÉLÉMENTS

DE LA

THÉORIE DES PROBABILITÉS

ÉLÉMENTS

DE LA

THÉORIE DES PROBABILITÉS

DU MÊME AUTEUR

Librairie GAUTHIER-VILLARS

Leçons sur la Théorie des Fonctions, 1898.
Leçons sur les Fonctions entières, 1900.
Leçons sur les Séries divergentes, 1901.
Leçons sur les Séries à termes positifs, 1902.
Leçons sur les Fonctions méromorphes, 1903.
Leçons sur les Fonctions de variable réelle, 1905.
Leçons sur la Théorie de la Croissance, 1909.

Librairie VUIBERT et NONY

Introduction à l'étude de la Théorie des nombres et de l'Algèbre supérieure (en collaboration avec M. Jules DRACH), 1895.

Librairie ARMAND COLIN

Cours élémentaire de Mathématiques. (*Arithmétique, Algèbre, Géométrie, Trigonométrie*), 1901-1905.

Librairie TEUBNER, à Leipzig

Die Elemente der Mathematik (traduction du Cours élémentaire précédent, par P. STÆCKEL), 1908 et 1909.

SAINT-AMAND (CHER). — IMPRIMERIE BUSSIÈRE.

COURS DE LA FACULTÉ DES SCIENCES DE PARIS

ÉLÉMENTS

DE LA

THÉORIE DES PROBABILITÉS

PAR

ÉMILE BOREL

PROFESSEUR-ADJOINT À LA FACULTÉ DES SCIENCES DE PARIS

PROBABILITÉS DISCONTINUES — PROBABILITÉS CONTINUES — PROBABILITÉS DES CAUSES

PARIS

LIBRAIRIE SCIENTIFIQUE A. HERMANN & FILS

LIBRAIRES DE S. M. LE ROI DE SUÈDE

6, RUE DE LA SORBONNE, 6

1909

06980

PRÉFACE

La *Théorie des Probabilités*, qu'on appelle aussi *Calcul des Probabilités* est utilisée de plus en plus dans de nombreuses questions de physique, de biologie, de sciences économiques. Ceux qui s'intéressent a ces applications n'ont pas toujours les loisirs d'étudier à fond les théories mathématiques qui se rattachent aux probabilités : ces théories n'ont d'ailleurs pour eux qu'un médiocre intérêt ; ce qui leur importe surtout c'est, avec avec la connaissance des résultats essentiels, celle des méthodes générales par lesquelles ces résultats sont obtenus ; il est évidemment nécessaire d'avoir réfléchi sur ces méthodes pour pouvoir appliquer avec sûreté les résultats bruts du calcul à des questions concrètes.

C'est à ce point de vue que j'ai écrit ces *Éléments* : je n'ai pas craint d'insister longuement sur les problèmes les plus simples, dans lesquels le mécanisme du calcul ne dissimule pas la méthode suivie. Si je n'ai point omis certains développements mathématiques qui peuvent intéresser quelques lecteurs, ces développements occupent peu de place et ne sont jamais indispensables à la compréhension de l'Ouvrage : celui-ci peut être lu d'un bout à l'autre par un lecteur connaissant simplement la définition de l'intégrale définie et les notions d'algèbre et de géométrie que cette définition suppose [1].

[1] Ce lecteur pourra omettre les nos 19 et 28.

Mais, si j'ai tenu à rester élémentaire, je me suis efforcé d'éliminer les développements de science amusante : les problèmes empruntés aux jeux de hasard ont été choisis uniquement pour illustrer une théorie générale. Il m'a été ainsi possible, en éliminant tout le superflu, de donner les principes essentiels de la théorie dans un Ouvrage relativement peu étendu.

Dans le Livre I j'étudie les *probabilités discontinues*, en insistant tout particulièrement sur le type le plus simple : les problèmes posés par *le jeu de pile ou face*. La véritable signification de la *loi des grands nombres* me paraît être mise ainsi en évidence de la manière à la fois la plus claire et la plus élémentaire.

Le Livre II est consacré aux *probabilités continues* ou *probabilités géométriques* ; c'est à cette catégorie de probabilités que se rattachent les plus importantes théories de la physique moderne, en particulier *la théorie cinétique des gaz*, et le *principe d'irréversibilité* de la thermodynamique, sur lequel j'ai donné quelques brèves indications.

Enfin, il m'a paru bon de grouper dans le Livre III, les questions relatives à la *probabilité des causes*, en raison de l'importance particulière de cette théorie pour les applications. C'est à elle en effet que se rattachent la *théorie des erreurs d'observation*, la théorie des *probabilités statistiques*, les études *biométriques*, etc. Le cadre de cet Ouvrage ne comportait pas l'étude détaillée de ces diverses applications ; je me suis contenté de les passer brièvement en revue, en insistant sur les conditions dans lesquelles la théorie peut y être utilisée et sur la méthode à suivre pour aborder l'étude de chaque problème concret.

Pour la technique des méthodes spéciales de calcul à utiliser dans ces diverses applications, je ne puis que renvoyer aux Traités spéciaux ou aux Mémoires originaux.

Je serais heureux si cet ouvrage contribuait d'une part, à faire mieux connaître à ceux qui étudient les sciences expérimentales et économiques les principes d'une théorie dont la

connaissance leur est chaque jour plus nécessaire ; d'autre part, à convaincre quelques jeunes mathématiciens de l'importance des applications de la théorie des probabilités et a les encourager à s'y intéresser.

Émile BOREL

Février 1909.

LIVRE I

PROBABILITÉS DISCONTINUES

CHAPITRE I

LE JEU DE PILE OU FACE

1. Définition du jeu. — On jette en l'air une pièce de monnaie et l'on engage des paris sur le côté qui sera apparent après sa chute : l'un des côtés s'appelle pile et l'autre face. Tel est le problème le plus simple de probabilités, si l'on ajoute à l'hypothèse que les chances sont égales pour le côté pile et le côté face. Au sujet de cette égalité des chances, nous laisserons de côté toute discussion philosophique, la prenant simplement pour un fait expérimental ou, si l'on veut, pour la *définition* même du fait que la pièce employée est *bonne* pour ce jeu : nous admettons dès lors que l'on se sert d'une pièce bonne. Si quelque lecteur pense que le relief irrégulier de la pièce ne permettra jamais qu'elle soit rigoureusement *bonne*, il peut se figurer que l'on emploie un jeton dont la symétrie a été rigoureusement recherchée par une fabrication des plus soigneuses : cette symétrie parfaite [1] ne peut d'ailleurs être qu'un cas limite, une notion abstraite comme la notion de la ligne droite. Mais cette hypothèse peut être suffisamment réalisée dans la pratique pour pour que les conclusions auxquelles elle conduit s'accordent avec l'expérience, au degré d'approximation même des erreurs expérimentales : c'est tout ce que l'on peut demander à une science quelconque.

---

[1] A parler en toute rigueur, la symétrie absolue entrainerait l'indiscernabilité des deux côtés et par suite l'impossibilité du jeu ; si les deux côtés diffèrent en quoi que ce soit, ne serait-ce que par la couleur, on ne peut jamais être rigoureusement sûr que cette différence est sans influence.

2 Etude des cas simples. — Etudions d'abord les cas où l'on joue un petit nombre de parties et cherchons à voir quelles sont les circonstances qui pourront se produire. La première partie peut donner comme résultat pile ou face ; ces deux hypothèses sont également probables ; on convient de dire que la *probabilité* [1] de chacune d'elles est *un* sur *deux*, ce que nous écrirons 1 : 2 ou 0,5. Nous pouvons représenter ce résultat par le tableau suivant

$$\begin{cases} \text{P} & 1 : 2 \\ \text{F} & 1 : 2 \end{cases}$$

Quel qu'ait été le résultat de la première partie, la seconde partie peut donner comme résultat pile ou face, et chacune de ces hypothèses est également probable, car le résultat de la première partie est sans influence sur celui de la seconde : suivant l'heureuse expression de Joseph Bertrand, la pièce de monnaie n'a ni conscience ni mémoire. J'insiste un peu sur ce point, malgré son évidence, car si on ne le comprend point d'une manière parfaite, sans restriction ni réticence, il est inutile de poursuivre l'étude du calcul des probabilités.

C'est surtout l'habitude des jeux de hasard qui rend certains esprits réfractaires à cette notion de l'indépendance des événements successifs : comme ils ont observé que, dans une longue série, les coups de pile ou face sont à peu près également nombreux [2], ils en concluent qu'une longue série de coups ayant amené pile *doit* être suivie d'un coup face : c'est une *dette* que le jeu a contractée envers eux. Il suffit d'un peu de réflexion pour se convaincre à quel point cet anthropomorphisme est puéril : les raisons pour lesquelles les chances de pile et de face sont égales subsistent à chaque partie et l'on ne peut concevoir aucun mécanisme par lequel les résultats des parties antérieures pourraient modifier l'égalité des chances. Cette croyance anthropomorphique a la mémoire et a la cons-

---

[1] La définition précise de la probabilité sera donnée au chapitre suivant ; cette définition empirique nous suffira pour l'instant.

[2] L'observation des joueurs porte sur le rouge ou le noir à la roulette, ou sur le gain ou la perte du banquier au baccara, etc. ; nous laissons de côté ces jeux, parce qu'ils sont plus complexes que le jeu de pile ou face.

cience de la pièce de monnaie n'a donc aucun fondement positif : pour les esprits superstitieux, cela pourrait ne pas suffire à la condamner s'ils n'apercevaient pas d'autre moyen de rendre compte des observations qui donnent des nombres sensiblement égaux pour pile et pour face, sur un grand nombre de parties : mais nous allons précisément voir que l'on arrive à expliquer de la manière la plus satisfaisante ces résultats de l'expérience, en admettant l'indépendance des parties successives : tout prétexte disparaît donc pour contester cette indépendance.

Revenons donc à notre seconde partie : en combinant les résultats qu'elle peut donner avec les résultats de la première, nous obtenons le tableau suivant :

$$\begin{array}{l|l} \mathrm{P} & \mathrm{P} \\ & \mathrm{F} \\ \mathrm{F} & \mathrm{P} \\ & \mathrm{F} \end{array}$$

dans lequel la première colonne correspond à la première partie et la seconde colonne à la seconde partie : nous voyons que l'ensemble des deux parties conduit à quatre combinaisons possibles, toutes quatre également probables : on dira que la probabilité de chacune d'elles est *un quart*, c'est-à-dire $1 : 4$ ou $0{,}25$. On peut d'ailleurs les écrire comme il suit

$$\left.\begin{array}{ll} \mathrm{PP} & 1 : 4 \\ \mathrm{PF} & 1 : 4 \\ \mathrm{FP} & 1 : 4 \\ \mathrm{FF} & 1 : 4 \end{array}\right| \quad 2 : 4 = 1 : 2$$

Nous avons réuni par un trait les deux combinaisons PF et FP : elles ont en effet ceci de commun qu'elles renferment chacune une fois pile et une fois face ; *elles sont identiques quand on ne tient pas compte de l'ordre* : si l'on énonce le résultat brut des deux parties, on peut dire que trois hypothèses seulement sont possibles : ou les deux parties amènent pile, ou une seule amène pile, ou aucune n'amène pile : en d'autres termes, le joueur qui aurait joué pile peut gagner, ou 2 fois, ou 1 fois, ou 0 fois. Les probabilités respectives de ces hypothèses

ne sont pas égales, comme une absence totale de réflexion aurait pu le laisser croire : le cas où le joueur gagne 1 fois sur 2 parties peut se produire par les deux combinaisons différentes PF et FP : sa probabilité est donc *un demi*, tandis que la probabilité de chacun des autres cas est seulement *un quart*.

Supposons maintenant que l'on joue une troisième partie : un raisonnement analogue nous conduira au tableau suivant

$$\begin{array}{l|l|l}
P & P & P \\
 & & F \\
 & F & P \\
 & & F \\
F & P & P \\
 & & F \\
 & F & P \\
 & & F
\end{array}$$

dans lequel la première colonne correspond à la première partie, la seconde colonne à la seconde partie, la troisième colonne à la troisième partie : on voit qu'il y a en tout 8 combinaisons, dont chacune est aussi probable que les autres. Modifiant l'ordre du tableau précédent, nous intervertirons la quatrième et la cinquième lignes, pour rapprocher les combinaisons qui conduisent au même résultat global, abstraction faite de l'ordre des parties, et nous obtiendrons ainsi le tableau suivant

$$\begin{array}{lll}
PPP & 1:8 & \\
PPF & 1:8 & \\
PFP & 1:8 & 3:8 \\
FPP & 1:8 & \\
PFF & 1:8 & \\
FPF & 1:8 & 3:8 \\
FFP & 1:8 & \\
FFF & 1:8 &
\end{array}$$

On voit que le joueur qui joue pile, par exemple, a 1 chance sur 8 de gagner 3 parties, 3 chances sur 8 d'en gagner 2, 3 chances sur 8 d'en gagner 1, et 1 chance sur 8 d'en gagner 0.

On formerait de la même manière les tableaux relatifs au cas de quatre parties ; nous donnerons seulement le second des tableaux, où, pour la commodité de la lecture, nous remplaçons la lettre F par un point.

4P	PPPP	1 : 16	
3P	PPP.	1 : 16	4 : 16 = 1 : 4
	PP.P	1 : 16	
	P.PP	1 : 16	
	.PPP	1 : 16	
2P	PP..	1 : 16	6 : 16 = 3 : 8
	P.P.	1 : 16	
	P..P	1 : 16	
	.PP.	1 : 16	
	.P.P	1 : 16	
	..PP	1 : 16	
1P	P...	1 : 16	4 : 16 = 1 : 4
	.P..	1 : 16	
	..P.	1 : 16	
	...P	1 : 16	
0P	....	1 : 16	

On voit que le joueur qui joue pile a 1 chance sur 16 de gagner les 4 parties, 4 chances sur 16 ou 1 sur 4 d'en gagner 3, 6 chances sur 16 ou 3 sur 8 d'en gagner 2, 4 chances sur 16 ou 1 sur 4 d'en gagner 1, 1 chance sur 16 de n'en gagner aucune.

3. Remarques sur les cas simples. — Sans qu'il soit nécessaire de pousser plus loin cette étude, on voit par quel mécanisme simple, on arrive en partant de probabilités supposées égales pour chaque partie, à trouver des probabilités inégales pour les divers résultats globaux possibles d'un ensemble de parties. Par exemple, dans le cas de 4 parties les combinaisons dans lesquelles le résultat global est 2 fois pile et 2 fois face sont au nombre de 6 ; chacune d'elles, prise individuellement, n'est pas plus probable que la combinaison qui amène pile 4 fois de suite ; mais, lorsque l'on considère leur ensemble, il y

a 6 fois plus de chances pour obtenir l'une d'elles, non désignée à l'avance, que pour obtenir la combinaison unique PPPP. Si deux joueurs font fréquemment des séries de 4 parties à pile ou face et s'ils observent avec soin les résultats du jeu, ils constateront sans peine que les cas les plus fréquents, au point de vue du résultat global, sont ceux où chacun d'eux gagne deux parties ; ces cas devront se produire en moyenne [1] 6 fois sur 16 ; les cas où l'un d'eux gagne 3 parties se produiront 4 fois sur 16 pour chacun d'eux et par suite 8 fois sur 16 si l'on ne spécifie pas lequel des deux gagne les trois parties : ils sont alors plus nombreux que les cas où l'égalité se produit ; enfin il arrivera seulement une fois sur 16 que l'un d'eux gagnera les 4 parties et par suite 2 fois sur 16 que les 4 parties seront gagnées par le même joueur, non désigné d'avance). Mais ce serait une très grave erreur de conclure de ces constatations que la combinaison PPPF, par exemple, est plus probable que la combinaison PPPP, c'est-à-dire que, les trois premières parties ayant donné pile, la quatrième a une tendance particulière à donner face. Si les combinaisons donnant pile 3 fois sont plus probables que la combinaison donnant pile 4 fois, ce n'est pas parce que chacune d'elles est plus probable, mais uniquement parce qu'elles sont plus nombreuses : il y en a 4, car la partie unique donnant face peut être soit la première, soit la seconde, soit la troisième, soit la quatrième ; c'est seulement ce dernier cas qui peut se produire lorsque les trois premières parties ont donné pile.

4. Triangle arithmétique de Pascal. — Lorsque l'on considère un grand nombre de parties successives de pile ou face, il est aisé de donner une règle faisant connaître la probabilité d'amener m fois pile sur n parties. Il est clair que le nombre total des combinaisons possibles est 2^n, et que chacune de ces combinaisons est également probable car à chacune des combinaisons possibles pour les $n-1$ premières parties en corres-

[1] Nous anticipons un peu, par cette affirmation, sur les résultats du Chapitre IV. Aussi est-il bon de faire observer que cette affirmation est donnée à titre d'exemple et qu'il n'y a pas de cercle vicieux dans les raisonnements.

pondent 2 pour les n parties, puisque la $n^{\text{ième}}$ partie peut donner soit pile, soit face ; c'est ainsi que nous avons trouvé 2 combinaisons dans le cas d'une partie, 4 dans le cas de deux parties, 8 dans le cas de trois parties, etc.

Parmi ces 2^n combinaisons, combien y en a-t-il renfermant m fois pile ? Nous désignerons leur nombre par C_n^m et nous remarquerons que, pour obtenir m fois pile, il faut ou bien que les $n-1$ premières épreuves aient donné m fois pile et que la $n^{\text{ième}}$ donne face, ou bien que les $n-1$ premières épreuves aient donné $m-1$ fois pile et que la $n^{\text{ième}}$ donne pile. La première alternative fournit autant de combinaisons qu'il en existe donnant m fois pile sur $n-1$ épreuves, c'est-à-dire C_{n-1}^m combinaisons ; de même, la seconde alternative fournit C_{n-1}^{m-1} combinaisons on a donc

$$C_n^m = C_{n-1}^{m-1} + C_{n-1}^m .$$

On déduit de cette formule la règle suivante, qui conduit au triangle arithmétique de Pascal (1) :

On inscrit sur une première ligne le nombre 1 *répété deux fois ; puis on calcule chaque nombre à inscrire dans les lignes suivantes en ajoutant au nombre inscrit immédiatement au-dessus de lui le nombre inscrit à la gauche de celui auquel on l'ajoute ; dans l'application de cette règle, on suppose mentalement que les lignes sont prolongées à droite et à gauche par des zéros.* On obtient ainsi le tableau suivant :

1	1									
1	2	1								
1	3	3	1							
1	4	6	4	1						
1	5	10	10	5	1					
1	6	15	20	15	6	1				
1	7	21	35	35	21	7	1			
1	8	28	56	70	56	28	8	1		
1	9	36	84	126	126	84	36	9	1	
1	10	45	120	210	252	210	120	45	10	1

(1) Le *Traité du triangle arithmétique* est reproduit (texte latin et traduction française) dans le Tome III des œuvres de Blaise Pascal, publiées par Léon Brunschvicg et Pierre Boutroux (Collection *les grands écrivains de la France*).

qu'il serait possible de prolonger et qui est de la plus grande utilité pour la solution des problèmes simples de probabilités. Les 4 premières lignes renferment les nombres que nous avons déjà trouvés au n° 2 dans les cas de 1, 2, 3, 4 parties ; la 5ième ligne s'obtient par le calcul suivant

$$\begin{aligned} 0 + 1 &= 1 \\ 1 + 4 &= 5 \\ 4 + 6 &= 10 \\ 6 + 4 &= 10 \\ 4 + 1 &= 5 \\ 1 + 0 &= 1. \end{aligned}$$

La somme des nombres de cette cinquième ligne est, d'après ces égalités même, double de la somme des nombres de la quatrième : elle est effectivement égale à $32 = 2 \times 16$. Il y a donc 32 combinaisons possibles pour 5 parties, parmi lesquelles 1 donnant 5 fois pile, 5 donnant 4 fois pile, 10 donnant 3 fois pile, 10 donnant 2 fois pile, 5 donnant 1 fois pile, 1 donnant 0 fois pile.

On peut remarquer que les diverses lignes du tableau sont symétriques, c'est-à-dire que les nombres équidistants des extrêmes sont égaux. Dans les lignes de rang pair, correspondant à un nombre pair de parties, les nombres vont en croissant jusqu'au milieu : le nombre le plus élevé correspond à un partage égal des résultats entre pile et face : dans les lignes de rang impair, ce partage égal ne peut être réalisé et il y a au milieu de la ligne deux nombres égaux, plus grands que les autres et correspondant aux deux cas où l'un des deux côtés de la pièce apparaît une fois de plus que l'autre.

5. Quelques problèmes élémentaires.

Problème I. *Un joueur à pile ou face a-t-il plus de chances de gagner* 3 *parties sur* 4 *ou d'en gagner* 5 *sur* 8 ? Il convient de préciser cet énoncé en indiquant si le joueur doit gagner précisément 3 parties et 5 parties ou s'il doit gagner *au moins* 3 parties ou 5 parties ; en d'autres termes, si le joueur gagne 6 parties, doit-on considérer qu'il en a gagné 5 ? C'est évidem-

ment la une question de convention ; nous examinerons successivement les deux cas.

Premier cas : l'énoncé est pris au sens strict, c'est-à-dire que les nombres de parties gagnées doivent être précisément 3 et 5. On voit immédiatement sur le triangle arithmétique que, pour gagner 3 parties sur 4 il y a 4 cas favorables sur 16 et que pour gagner 5 parties sur 8, il y a 56 cas favorables sur 256 ; dans la première hypothèse la probabilité de gagner est donc égale à $\frac{4}{16}$ c'est-à-dire à $\frac{1}{4}$, dans la seconde hypothèse elle est $\frac{56}{256}=\frac{7}{32}$; elle est donc moins élevée ; *on a plus d'avantage à parier que l'on gagnera* 3 *parties sur* 4 *qu'à parier que l'on gagnera* 5 *parties sur* 8.

Second cas : l'énoncé est pris au sens large, c'est-à-dire que les nombres de parties gagnées doivent être au moins égaux à 3 et à 5. Le triangle arithmétique montre que pour gagner 4 ou 3 parties sur 4, il y a un nombre de cas favorables égal à $1+4=5$ sur 16 ; pour gagner 8, 7, 6 ou 5 parties sur 8, il y a un nombre de cas favorables égal à $1+8+28+56=93$ sur 256 : dans la première hypothèse, la probabilité de gagner est $\frac{5}{16}$ ou $\frac{80}{256}$; dans la seconde hypothèse elle est $\frac{93}{256}$, c'est à-dire plus élevée ; *on a plus d'avantage à parier que l'on gagnera au moins* 5 *parties sur* 8 *qu'à parier qu'on en gagnera au moins* 3 *sur* 4.

Problème II. *Y a-t-il plus d'avantage à parier que l'on gagnera au moins* 5 *parties sur* 7 *ou que l'on en gagnera au moins* 6 *sur* 9 *?*

Dans la première hypothèse le nombre des cas favorables est

$$1+7+21=29$$

sur 128 : la probabilité de gain est

$$\frac{29}{128}=\frac{1}{4}+\frac{3}{128}.$$

Dans la seconde hypothèse, le nombre des cas favorables est

$$1+9+36+84=130$$

sur 5×2 ; la probabilité de gain est

$$\frac{150}{512} = \frac{1}{4} + \frac{1}{256}.$$

Elle est plus élevée : c'est donc cette seconde hypothèse qui est plus avantageuse.

Il est inutile de multiplier les exemples de ces problèmes simples : pour traiter des cas un peu plus compliqués, il est nécessaire d'introduire la notion d'espérance mathématique.

6 Espérance mathématique. — On appelle *espérance mathématique* d'un joueur le produit de son gain possible par la probabilité qu'il a de le réaliser. Nous nous plaçons tout d'abord dans le cas le plus simple où le gain possible est unique.

Par exemple Pierre doit recevoir 100 francs dans le cas où une partie de pile ou face donnera pile : la probabilité est 0,5 : son espérance mathématique est donc 50 francs. Paul doit recevoir 1 000 francs dans le cas où deux parties consécutives donneront toutes deux pile : son espérance mathématique est donc

$$1\,000 \times 0{,}25 = 250 \text{ francs.}$$

On voit que l'expression *espérance mathématique* doit être regardée comme un vocable unique, ayant un sens bien déterminé qu'on ne doit pas chercher à interpréter d'après le sens usuel des deux termes : *espérance* et *mathématique*. Au sens vulgaire du mot, Paul a une certaine espérance de toucher 1 000 francs : si cette espérance ne se réalise pas, il ne touchera rien du tout ; en aucun cas, il ne touchera 250 francs, montant de son *espérance mathématique*. Ceci étant bien entendu, il n'y a pas d'inconvénient à conserver cette expression, qui est consacrée par l'usage : la création de termes nouveaux présente en effet rarement de réels avantages.

On dit qu'un jeu est *équitable* lorsque l'espérance mathématique du joueur est égale à sa mise : l'espérance mathématique d'une certaine somme peut donc être échangée contre cette somme, *dans le cas où l'on trouve un joueur disposé à accepter*

un certain jeu équitable. Par exemple, si je dois toucher 2 francs au cas où une partie de pile ou face donnera pile, mon espérance mathématique est 1 franc et peut être vendue ce prix à un joueur disposé à prendre ma place. Dans ce cas là, je puis espérer trouver facilement à *vendre* mon espérance mathématique ; il n'en serait pas de même si l'enjeu de la partie de pile ou face était de cent millions : dans certains cas la valeur *commerciale* d'une espérance mathématique est donc inférieure à sa valeur numérique ; dans certains cas, au contraire, la valeur commerciale de l'espérance mathématique est supérieure à sa valeur numérique : c'est une conséquence du goût du public pour les loteries et des restrictions mises par la loi à leur développement : un billet de loterie auquel est attaché une espérance mathématique de trente centimes [1] trouve facilement preneur à un franc.

Le grand avantage de la notion de l'espérance mathématique est le suivant : la combinaison de probabilités diverses entraîne à des calculs parfois compliqués, comme nous le verrons dans le chapitre suivant ; au contraire, pour les espérances mathématiques, la règle est simple et intuitive : *pour avoir l'espérance mathématique totale qui s'attache à plusieurs événements fortuits, il suffit de faire la somme des espérances mathématiques correspondant à chacun d'eux*.

Paul doit jouer trois parties de pile ou face ; s'il gagne la première il touchera 10 francs ; s'il gagne la seconde, 20 francs ; s'il gagne la troisième, 50 francs. Son espérance mathématique totale est $5 + 10 + 20 = 35$ francs ; telle est la somme pour laquelle il pourrait vendre ses diverses chances à un joueur disposé à jouer un jeu équitable. Si ce fait ne paraît pas évident, il suffit d'observer que si un même acheteur achète l'espérance mathématique de Paul et celle de son partenaire, il doit les payer toutes deux le même prix, puisque la situation de Paul et celle de son partenaire sont identiques : or cet acheteur est assuré de toucher 70 francs, quoiqu'il arrive : donc, pour que

[1] Pour trouver l'espérance mathématique attachée à un billet de loterie, on n'a qu'à diviser la somme totale des lots par le nombre des billets ; ce serait la valeur que devrait avoir le billet si l'entrepreneur de la loterie ne faisait aucun gain et n'avait pas de frais de publicité.

le jeu ne lui apporte ni bénéfice assuré, ni perte certaine, il doit payer 35 francs à chacun des deux partenaires.

On doit admettre comme évident que si un jeu se compose de plusieurs parties successives et si chaque partie est équitable, le jeu est équitable dans son ensemble : ce principe est d'une grande utilité pour la solution de certaines questions.

7. Problèmes sur l'espérance mathématique. — Problème III. *Trois joueurs*, A, B, C, *jouent à pile ou face dans les conditions suivantes :* A *et* B *font une première partie, à la suite de laquelle le perdant se retire et cède sa place à* C ; *il en est de même après chaque partie : le perdant se retire pour céder sa place au troisième joueur ; le joueur qui aura gagné deux parties consécutives touchera une somme* m ; *quelle est l'espérance mathématique de chaque joueur :* 1° *après la première partie ?* 2° *au début du jeu ?*

Supposons que A gagne la première partie et désignons, après cette première partie, par a, b, c les espérances mathématiques de A, B, C qui sont respectivement le joueur qui reste, celui qui sort et celui qui rentre. Si A gagne la seconde partie, le jeu sera fini et A en sera le gagnant : si A la perd, il deviendra le joueur qui sort, B sera celui qui rentre et C celui qui reste. Dans la première hypothèse (A gagnant), A touche m, B et C ne touchent rien ; dans la seconde (A perdant) l'espérance mathématique de A devient b, celle de B devient c, et celle de C devient a. Nous écrirons que l'espérance mathématique de chacun des joueurs est égale à la somme des espérances mathématiques qui résultent de chacune de ces deux hypothèses ; comme la probabilité de chaque hypothèse est *un demi* nous obtenons ainsi les équations

$$a = \frac{m}{2} + \frac{b}{2}$$

$$b = 0 + \frac{c}{2}$$

$$c = 0 + \frac{a}{2}$$

qui sont faciles a résoudre et donnent

$$a = \frac{4}{7}m\,; \qquad b = \frac{1}{7}m\,; \qquad c = \frac{2}{7}m$$

Telles sont les espérances mathématiques apres la premiere partie : pour obtenir les espérances mathematiques a', b', c' avant cette premiere partie, il suffit d'observer que l'espérance mathématique de C ne peut pas être modifiée par le résultat de cette premiere partie ; d'autre part *la somme* des espérances mathématiques de A et de B n'est pas non plus modifiée ; car, pour celui qui aurait acheté à A et à B leurs chances de gain, il est indifférent que l'un ou l'autre gagne cette premiere partie : enfin, on a forcément $a' = b'$; on obtient ainsi finalement :

$$a' = b' = \frac{a+b}{2} = \frac{5}{14}m \qquad c' = c = \frac{2}{7} = \frac{4}{14}m$$

On voit qu'au début du jeu, la situation des joueurs A et B est un peu meilleure que celle du joueur C : si la somme à gagner m est $1\frac{4}{7}$ francs, l'espérance mathématique de A ou de B vaut 5 francs, et celle de C vaut seulement 4 francs.

Nous reviendrons sur ce probleme, que l'on appelle parfois *problème de la poule*, pour l'étudier a l'aide des théoremes sur les probabilités.

Problème IV. *Pierre joue avec Paul n parties successives de pile ou face et l'on s'engage à lui verser a francs chaque fois que pile sera sorti 2 fois sur 3 parties consécutives et b francs lorsque pile sera sorti 3 fois de suite. Quelle est l'espérance mathématique de Paul ?* Il suffit d'observer qu'il y a $n-2$ groupes de 3 parties consécutives, commençant respectivement par la premiere, la seconde, ... la $n-2^{\text{ième}}$. Pour chacun de ces groupes l'espérance mathématique de Paul est

$$\frac{3a}{8} + \frac{b}{8}$$

puisqu'il a 3 chances sur 8 de gagner a et 1 chance sur 8 de gagner b : l'espérance mathématique totale est donc

$$\frac{(n-2)(3a+b)}{8}$$

Nous donnerons plus loin des types de problèmes plus compliqués.

8 Formules générales pour n parties. — Dans le cas où le nombre n de parties de pile ou face est assez grand, l'emploi du triangle arithmétique cesse d'être commode ; il est préférable d'utiliser une formule faisant directement connaître le nombre des combinaisons fournissant q fois pile sur n parties. La théorie des combinaisons, dans laquelle intervient aussi le triangle arithmétique, nous apprend que ce nombre est donné par la formule [1]

$$C_n^q = \frac{n(n-1)\dots(n-q+1)}{1.2\dots q} = \frac{n!}{q!\,(n-q)!}$$

Mais il est aisé de démontrer cette formule par des considérations directes. Nous nous proposons donc de rechercher combien de combinaisons différentes peuvent amener q fois pile sur n parties consécutives : si dans l'une de ces combinaisons, nous remplaçons l'une des q parties ayant amené pile par une partie ayant amené face, nous obtenons l'une des C_n^{q-1} combinaisons qui amènent pile $q-1$ fois : comme nous pouvons faire cette substitution de q manières différentes sur chacune des C_n^q combinaisons puisque chacune d'elles renferme pile q fois, nous obtenons en tout qC_n^q combinaisons renfermant $q-1$ fois pile et $n-q+1$ fois face. Mais il est clair que chacune des C_n^{q-1} combinaisons renfermant $q-1$ fois pile et $n-q+1$ fois face est ainsi obtenue $n-q+1$ fois, puisque, en remplaçant dans une telle combinaison l'une des $n-q+1$ *face* par *pile*, on trouve une des combinaisons primitives, d'où elle se déduit par la substitution inverse ; on a donc

$$qC_n^q = (n-q+1)\,C_n^{q-1}$$

c'est-à-dire

$$C_n^q = \frac{n-q+1}{q}\,C_n^{q-1}$$

---

[1] Nous posons, suivant l'usage, $n! = 1.2.3\dots n$, ce produit s'énonce *factorielle de n*.

On aurait de même

$$C_n^{q+1} = \frac{n-q+2}{q+1} C_n^{q-1}$$

$$C_n^{q+2} = \frac{n-q+3}{q+2} C_n^{q+1}$$

$$\cdots\cdots\cdots\cdots$$

$$C_n^2 = \frac{n-1}{2} C_n^1$$

$$C_n^1 = n$$

cette dernière équation résultant du fait que la partie unique amenant pile peut occuper soit le premier, soit le second, ... soit le $n^{\text{ième}}$ rang, ce qui fait n combinaisons distinctes. En multipliant ces diverses équations membre à membre on retrouve bien la formule donnée plus haut.

On peut observer enfin, et c'est là la démonstration la plus utile pour la suite, que si l'on considère le produit

$$(P+F)^n = (P+F)(P+F)(P+F)\ldots\ldots(P+F)$$

et si on effectue les opérations indiquées dans le second membre en commençant par la droite et ayant soin de conserver l'ordre des facteurs des produits partiels dans l'ordre dans lequel ils se présentent, sans effectuer la réduction des termes semblables, le résultat obtenu consiste précisément en la somme des 2^n combinaisons possibles des deux lettres P et F. Si l'on fait ensuite la réduction des termes semblables, il est clair que le coefficient d'un terme tel que $P^q F^{n-q}$ est précisément égal au nombre C_n^q des combinaisons renfermant q fois la lettre P. On retrouve ainsi la formule du binome de Newton

$$(P+F)^n = P^n + C_n^1 P^{n-1}F + \ldots + C_n^q P^{n-q}F^q + \ldots + F^n$$

dans laquelle on a $C_n^q = C_n^{n-q}$. Les coefficients C_n^q reçoivent souvent le nom de *coefficients binomiaux* ; ils jouent un rôle très important dans les questions de probabilité ; on vient de voir par quel mécanisme simple ils s'y introduisent.

Remarquons qu'il résulte des considérations précédentes que l'on a

$$1 + C_n^1 + C_n^2 + \ldots + C_n^q + \ldots + C_n^{n-1} + 1 = 2^n$$

car le nombre total des combinaisons possibles est 2^n : cette égalité s'obtient d'ailleurs aussi en développant $(1+1)^n$ par la formule du binome de Newton.

Nous donnerons plus loin (n°s 17 et 19) des formules pour le calcul approché des valeurs de $n!$ et par suite des coefficients C_n^p ; ces formules approchées sont d'autant plus avantageuses que n est plus grand : lorsque n ne dépasse pas 30, on utilisera avec profit une table numérique faisant connaître les logarithmes vulgaires des valeurs de $n!$: on trouve cette table dans la plupart des tables de logarithmes.

9. **Remarques sur quelques paradoxes.** — Nous pourrions borner là l'exposé des principes essentiels de la théorie du jeu de pile ou face : ces principes étant bien établis, les conséquences que nous en déduirons par des raisonnements purement logiques sont rigoureusement démontrées, et par suite toute assertion contraire à ces conséquences devra être regardée comme inexacte, sans qu'il soit nécessaire d'examiner les arguments sur lesquels on prétend la baser. Cette maniere de procéder est la plus conforme à l'esprit mathématique : il me paraît cependant préférable de ne pas m'y tenir, car tout le monde n'a pas l'esprit mathématique et, en ce qui concerne les questions de probabilité, beaucoup d'esprits, excellents par ailleurs, ont une certaine méfiance des raisonnements logiques et sont disposés à leur préférer des raisons de sentiment. J'ai eu récemment l'occasion de constater cette tendance chez un des esprits les plus distingués de notre temps, bien connu par ses publications scientifiques et philosophiques, et dont l'éducation mathématique a été très sérieuse. Il m'a dès lors semblé qu'il valait mieux ne pas traiter ces tendances par le pur dédain que serait en droit de leur opposer un mathématicien qui jugerait entièrement superflu de convaincre, du moment que ses raisonnements sont irréprochables. Il y a, en effet, à mon avis, un très grand intérêt scientifique et social à ce que les principes fondamentaux du calcul des probabilités soient admis sans restriction par le plus de personnes possible : si donc quelques arguments peuvent amener ce résultat, il vaut la peine d'y con-

sacrer quelques lignes bien que, au point de vue mathématique absolu, ils soient inutiles.

L'une des principales sources de ces raisonnements paradoxaux sur lesquels nous aurons à revenir (voir n° **18**) est la suivante : *on considère un événement futur comme réalisé, sous prétexte que l'expérience a prouvé qu'il est extrêmement probable.* On commet ainsi une erreur, sans doute très petite, mais l'accumulation répétée de telles erreurs suffit pour conduire à des conséquences entièrement inexactes. Nous étudierons plus loin le mécanisme mathématique de cette accumulation d'erreurs ; pour l'instant, nous nous contenterons de montrer par l'absurde le défaut de ce genre de raisonnements.

Supposons que l'on joue un très grand nombre de parties successives à pile ou face et que l'on note tous les résultats ; on saura à chaque instant quel est le gain ou la perte d'un joueur qui aurait toujours parié pour pile, l'enjeu étant toujours le même à chaque partie. Si l'on fait effectivement l'expérience, on constate aisément qu'au bout d'un certain nombre de parties, souvent petit, dépassant rarement 100 et presque jamais 1 000 le gain et la perte se trouvent réduits à zéro : lorsqu'on en est arrivé là, il y a une chance sur deux pour que la partie suivante amène un gain ; s'il en est ainsi, nous dirons que l'ensemble des parties jouées constitue une bonne série ; sinon, nous continuerons le jeu jusqu'à un nouveau retour à zéro et il y aura de nouveau une chance sur deux pour que la partie suivante soit gagnée : s'il en est ainsi, nous aurons une bonne série ; sinon nous continuerons encore le jeu et nous finirons bien par arriver à obtenir *une bonne série*, puisque nous avons une chance sur deux d'y parvenir chaque fois que nous revenons au zéro. Pratiquement, l'expérience montrera à celui des lecteurs qui voudra la tenter que l'on obtient généralement une bonne série après un petit nombre de coups et que l'on y parvient *sûrement* [1]) pourvu que l'on ait la patience de jouer, s'il est nécessaire, quelques milliers de coups.

---

(1) Je souligne ce mot, car c'est par lui que s'introduit l'erreur de raisonnement ; il faudrait dire *presque sûrement*, c'est-à-dire que la probabilité pour ne pas réussir est extrêmement faible (voir au n° **18** un calcul détaillé) ; il y a donc certitude pratique, et c'est ce que montre l'expérience, qui réussit

Ceci étant admis, Paul joue avec Pierre à pile ou face et est décidé à prolonger le jeu jusqu'à ce qu'il ait réalisé une bonne série [1] ; il réalise ainsi un gain égal à la mise. Il cesse alors le jeu qu'il reprend le lendemain contre Jean : il le prolonge de même jusqu'à ce qu'il ait réalisé une bonne série : il peut ainsi continuer tous les jours et gagner régulièrement une somme égale à la mise. Si nous supposons maintenant que Paul joue toujours contre le même adversaire Pierre une suite indéfinie de parties, Paul peut sans attendre au lendemain, considérer son jeu comme interrompu après chaque bonne série et recommencer à noter à ce moment là ; comme rien ne distingue un instant d'un autre et que l'on peut toujours supposer que le jeu commence à un instant quelconque, Paul réalisera ainsi un nombre illimité de bonnes séries successives et par suite, un gain illimité ou, du moins, qui ne sera limité que par la lenteur du jeu et la durée de la vie humaine. Mais, sur la même succession de parties, Pierre peut faire le même raisonnement ; son gain est donc aussi illimité, à condition que l'on puisse jouer assez longtemps : telle est la conséquence absurde à laquelle on aboutit : chacun des joueurs réalise un gain qui croît proportionnellement au temps. Nous nous contenterons pour l'instant, d'avoir mis en garde le lecteur contre les inconvénients de certains modes de raisonnement ; nous étudierons d'une manière plus approfondie cette question lorsque nous aurons acquis les principes nécessaires.

---

toujours (c'est-à-dire qui n'échouera à peine qu'une fois en 100 ans si l'on joue une partie par seconde et que l'on note tous les résultats) ; mais ce n'est pas une certitude mathématique.

[1] Si la première partie est gagnée, elle constitue à elle seule une bonne série.

CHAPITRE II

QUELQUES DÉFINITIONS ET QUELQUES THÉORÈMES

10. Définition de la probabilité. — On dit habituellement que *la probabilité est le rapport du nombre de cas favorables au nombre des cas possibles, lorsque tous les cas sont regardés comme également probables.* Cette définition renferme en apparence un cercle vicieux : comment pourra-t-on savoir que tous les cas sont également probables, si l'on ne sait pas ce que c'est que la probabilité ? En réalité, il n'y a pas de cercle vicieux à supposer que l'on a la notion vulgaire du sens des mots « également probable », lorsque l'on veut définir le sens mathématique précis du mot probabilité. Les logiciens qui prétendent construire des systèmes entièrement logiques, sans cercle vicieux, oublient qu'il est impossible de ne pas utiliser le langage usuel, ne serait-ce que pour définir les termes scientifiques que l'on emploie et pour construire les phrases dont on se sert : or, le langage usuel doit être considéré comme une acquisition globale de chaque individu, acquisition qui suppose un grand nombre de cercles vicieux.

Il est cependant naturel de se demander dans quels cas on regardera divers événements possibles comme également probables. La meilleure réponse à cette question sera donnée par quelques exemples : si l'on jette un dé ayant la forme d'un cube et dont la substance est homogène, il est également probable

qu'il montrera l'une quelconque de ses six faces : si, ayant mis dans une urne des boules de même forme et de même substance, ne différant que par la couleur et si, après avoir agité, on plonge la main dans l'urne, l'extraction de chacune des boules a la même probabilité. Il est d'autres cas plus compliqués, où l'égalité des probabilités sera postulée par définition, ou sera regardée comme un résultat de l'expérience : mais nous laisserons de côté ces questions, fort importantes au point de vue des applications, mais sans intérêt au point de vue de la théorie mathématique : pour que celle-ci soit légitime, il suffit qu'il y ait des cas où la définition est applicable [1] : la théorie s'applique à ces cas tout d'abord, et ensuite à tous les cas qui pourront se rencontrer et dans lesquels on aura constaté que la définition fondamentale s'applique aussi.

11. Propriétés essentielles de la probabilité. — Désignons par N le nombre total des cas possibles supposés tous également probables, bien entendu : nous ne reviendrons pas sur cette restriction et par n le nombre des cas favorables : la probabilité p est donnée par la formule

$$p = \frac{n}{N} ;$$

on voit que la probabilité est une fraction *toujours inférieure à l'unité* : elle devient égale à l'unité dans le cas seulement où $n = N$: tous les cas sont favorables : la probabilité se transforme en *certitude*.

En même temps que la probabilité d'un événement désiré, on a souvent intérêt à considérer la probabilité de l'événement contraire : nous supposons, pour l'instant, que tous les événements autres que l'événement favorables sont groupés sous cette même dénomination. Par exemple une urne renferme N boules parmi lesquelles n sont blanches, les autres étant rouges, noires, etc. On regarde comme l'événement favorable

[1] A parler strictement, on pourrait construire une théorie purement logique sans se préoccuper de l'existence possible de ses applications ; mais une telle théorie serait un pur jeu d'esprit sans intérêt, qui ne mériterait pas le nom de science.

l'extraction d'une boule blanche : l'événement contraire ou défavorable est alors l'extraction d'une boule qui ne soit pas blanche ; si nous désignons sa probabilité par q, nous avons

$$q = \frac{N - n}{N} = 1 - \frac{n}{N} = 1 - p$$

car le nombre des cas favorables à cet événement contraire est évidemment $N - n$, puisqu'il y a $N - n$ boules qui ne sont pas blanches.

La formule peut s'écrire

$$p + q = 1$$

c'est-à-dire que : *la somme des probabilités d'un événement et de l'événement contraire est égale à l'unité*. Il importe de bien remarquer que l'on entend ici par événement contraire tout ce qui n'est pas l'événement regardé comme favorable, on n'admet pas en ligne de compte diverses alternatives : par exemple, s'il s'agit pour un joueur de gagner une partie d'échecs, l'événement contraire comprend tous les cas où il ne la gagne pas : partie perdue et partie nulle.

Lorsqu'un événement est très probable, sa probabilité p est très voisine de l'unité et la probabilité q est par suite très voisine de zéro. On dit quelquefois, dans le langage ordinaire, que la probabilité est *très grande* ; cette manière de parler ne s'accorde pas avec la définition mathématique d'après laquelle p ne doit pas dépasser l'unité : au lieu de dire *probabilité très grande*, on devrait dire *probabilité très voisine de un*. Le langage usuel a son origine dans le fait suivant : en réalité, on compare l'éventualité de l'événement favorable à celle de l'événement contraire et lorsque la première éventualité dépasse de beaucoup la seconde, on dit que la probabilité est très grande : ce qui est très grand, *c'est le rapport de la probabilité de l'événement considéré à la probabilité contraire*. Ce rapport pourrait être appelé *probabilité relative* des deux événements, alors que la probabilité définie plus haut serait nommée la *probabilité absolue* : mais nous ne modifierons pas le langage usuel et dirons simplement *probabilité* et non *probabilité*

absolue, car c'est cette probabilité qui intervient presque toujours et il vaut mieux ne pas allonger inutilement l'expression qui la désigne. Il est toutefois bon de savoir que, dans le langage ordinaire, non scientifique, on emploie souvent le mot probabilité dans le sens de probabilité relative, c'est-à-dire que l'on pense au rapport R défini par la formule

$$R = \frac{p}{q} = \frac{p}{1-p} = \frac{n}{N-n}.$$

Par exemple, si une urne renferme 1 000 boules, dont 999 sont blanches et 1 seule noire, on a en regardant comme favorable l'extraction d'une boule blanche

$$\begin{aligned} p &= 0{,}999 \\ q &= 0{,}001 \\ R &= 999. \end{aligned}$$

La probabilité p de l'événement favorable est très voisine de 1 ; sa probabilité relative R est très grande.

Lorsque l'on envisage à la fois plusieurs éventualités qui s'excluent réciproquement et dont l'ensemble embrasse tous les cas possibles, la somme de leurs probabilités est égale à 1. Supposons, par exemple qu'une urne renferme N boules identiques, sauf que chacune porte inscrit l'un des nombres 1, 2, 3, ..., m : on désignera par n_1 le nombre des boules portant le nombre 1, par n_2 le nombre de celles qui portent le nombre 2, etc. La probabilité p_1 d'extraire l'une des boules qui portent le nombre 1 est évidemment

$$p_1 = \frac{n_1}{N}.$$

On a de même :

$$p_2 = \frac{n_2}{N}, \quad \dots, \quad p_m = \frac{n_m}{N}$$

et par suite :

$$p_1 + p_2 + \dots + p_m = \frac{n_1 + n_2 + \dots + n_m}{N} = 1$$

car la somme $n_1 + n_2 + \dots + n_m$ est égale au nombre total N des boules.

12. Probabilités totales. — Considérons une urne renfermant N boules, parmi lesquelles a sont rouges, b sont blanches, les autres n'étant ni blanches ni rouges. La probabilité α d'extraire une boule rouge est

$$\alpha = \frac{a}{N}.$$

La probabilité β d'extraire une boule blanche est

$$\beta = \frac{b}{N}.$$

La probabilité p d'extraire une blanche *ou* rouge est évidemment

$$p = \frac{a + b}{N}$$

et l'on aperçoit immédiatement la relation

$$p = \alpha + \beta$$

qui exprime le *théorème des probabilités totales*. Pour énoncer ce théorème, il importe de bien préciser les conditions dans lesquelles nous avons obtenu la relation précédente. Nous considérons comme événement favorable, dont nous recherchons la probabilité p, l'extraction d'une boule rouge *ou* blanche : cet événement favorable peut se produire de deux manières différentes, qui s'excluent réciproquement, car si la boule qu'on extrait est rouge, elle n'est pas blanche. Les probabilités respectives de ces deux éventualités sont α et β : *la probabilité p est égale à leur somme*. Il est visible que le raisonnement serait le même dans le cas où le nombre des éventualités favorables est supérieur à deux ; si, par exemple, on considérait comme favorable l'extraction d'une boule rouge, ou blanche, ou verte, ou violette, ou jaune, les autres couleurs étant exclues. Nous pouvons donc énoncer sous la forme suivante le théorème des probabilités totales.

Théorème. — *Lorsque l'événement dont on recherche la probabilité peut se produire de plusieurs manières différentes qui s'excluent réciproquement, la probabilité cherchée est égale à la somme des probabilités partielles correspondant à ces diverses manières.*

Dans les applications de ce théorème, il est très important de vérifier que la condition d'*exclusion réciproque* est bien vérifiée. Considérons, par exemple, le problème suivant.

Problème V. *Pierre et Paul jouent à pile ou face dans les conditions suivantes : si la première partie amène face, Pierre a gagné ; s'il n'en est pas ainsi, on joue deux autres parties, et, si sur l'ensemble des trois parties, face est amené au moins deux fois, Pierre a aussi gagné ; quelle est la probabilité qu'a Pierre de gagner ?*

On pourrait raisonner comme il suit : Pierre peut gagner de deux manières différentes, soit par le gain de la première partie, soit par le gain d'au moins deux parties sur trois. Or la probabilité de la première hypothèse est $\frac{1}{2}$, la probabilité de la seconde est aussi $\frac{1}{2}$; donc la probabilité totale est 1, c'est-à-dire la certitude. Cette conséquence est évidemment absurde : si l'on considère l'ensemble de trois parties de pile ou face, il est bien vrai que la probabilité d'amener face au moins deux fois est égale à $\frac{1}{2}$, car, sur 8 combinaisons possibles il y en a 4 favorables ;

FFP, FPF, PFF, FFF,

Mais, parmi ces quatre combinaisons, il y en a 3 qui ne peuvent se produire que si la première partie a donné face : or, dans ce cas là, Pierre gagne dès cette première partie et ne joue pas les deux autres ; ses chances de gain dans la seconde hypothèse sont donc exclues par les chances de gain dans la première. Il serait d'ailleurs aussi inexact de dire qu'il ne reste qu'une chance favorable sur 8, à savoir PFF et que par suite la proba-

bilité de la seconde hypothèse est $\frac{1}{8}$; car la moitié des 8 cas possibles se trouvent exclus par le fait que face n'a pas été amené par la première partie. Ce qu'il faut rechercher, c'est la probabilité de la seconde hypothèse face 2 fois sur 3 lorsque la première face a la première partie se trouve exclue; il faut donc amener face deux fois en deux parties la seconde et la troisième : la probabilité de cet événement est $\frac{1}{4}$; la probabilité qu'a Pierre de gagner dans les conditions indiquées est donc

$$\frac{1}{2} + \frac{1}{4} = \frac{3}{4}.$$

13. Probabilités composées. — Considérons deux urnes dont la première renferme N boules, parmi lesquelles a sont blanches et dont la seconde renferme N boules, parmi lesquelles a sont blanches. La probabilité α d'extraire une boule blanche de la première urne est

$$\alpha = \frac{a}{N}$$

et la probabilité α' d'extraire une boule blanche de la seconde urne est

$$\alpha' = \frac{a'}{N'}.$$

Supposons maintenant que l'on extraie une boule de chaque urne. Quelle est la probabilité p pour que ces boules soient blanches toutes les deux? Evaluons le nombre des cas possibles : pour cela supposons que nous ayons numéroté, afin de les distinguer entre elles, les boules de chaque urne : celles de la première sont numérotées de 1 à N et celles de la seconde de 1 à N'. L'extraction peut amener la boule 1 de la première urne et l'une quelconque des N' boules de la seconde; cela fait N' cas, tous également probables; on peut aussi amener la boule 2 de la première urne et l'une quelconque des N' boules de la seconde : cela fait encore N' cas, également probables entre eux, et de même probabilité que les précédents; à chacune des N boules de la première urne correspondent ainsi N' cas; il y a donc en

tout NN cas possibles (1). On compterait de même le nombre des cas favorables : à chacune des a boules blanches de la première urne correspondent a cas favorables, correspondant respectivement à l'extraction de chacune des a boules blanches de la seconde urne : il y a donc aa cas favorables et la probabilité p est donnée par la formule

$$p = \frac{aa}{NN}.$$

On voit que l'on a

$$p = \alpha\beta$$

et cette formule exprime le *théorème des probabilités composées*. Avant d'énoncer ce théorème, nous généraliserons les conditions dans lesquelles nous l'avons établi : supposons que nous ayons trois urnes, dont l'une sera toujours appelée la première et dont les deux autres, supposées peintes extérieurement, l'une en blanc, l'autre en noir, s'appelleront la *seconde urne blanche* et la *seconde urne noire*. On extrait d'abord une première boule de la première urne : si cette première boule extraite est blanche, on extrait la seconde boule de la *seconde urne blanche* : si la première boule extraite n'est pas blanche, on extrait la seconde boule de la *seconde urne noire*. Quelle est la probabilité pour avoir deux boules blanches ? Il est clair que si l'on désigne par α la probabilité d'extraire une boule blanche

---

(1) Si nous supposons : N = 6 et les N boules numérotées 0, 1, 2, 3, 4, 5 ; N = 10 et les N boules numérotées 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, les résultats des extractions pourront être représentés par le tableau suivant, où l'on a inscrit d'abord le numéro de la boule sortie de la première urne

00	10	20	30	40	50
01	11	21	31	41	51
02	12	22	32	42	52
03	13	23	33	43	53
04	14	24	34	44	54
05	15	25	35	45	55
06	16	26	36	46	56
07	17	27	37	47	57
08	18	28	38	48	58
09	19	29	39	49	59

Il est visible que les cas possibles sont au nombre de 60.

de la première urne et par β la probabilité d'extraire une boule blanche de la seconde urne *blanche*, on a encore

$$p = \alpha\beta.$$

En effet, on ne se sert de la seconde urne *noire* que dans le cas où le premier tirage n'a pas donné une boule blanche : donc, dans ce cas on ne se trouve certainement pas dans un cas favorable et le résultat du second tirage est sans importance : rien ne serait donc changé à la probabilité des cas favorables si l'on n'effectuait pas ce second tirage, ou bien si on l'effectuait dans une urne de composition absolument quelconque, *par exemple dans la seconde urne blanche* ; nous nous trouvons ainsi ramenés au cas déjà traité, mais ces remarques permettent de donner un énoncé plus général.

Théorème. — *Lorsque l'évènement dont on recherche la probabilité consiste dans la production successive de deux évènements, la probabilité cherchée est égale au produit de la probabilité du premier de ces évènements par la probabilité pour que le second se produise lorsque le premier s'est produit. Plus généralement, si la production successive de plusieurs évènements est nécessaire, on doit multiplier les probabilités diverses de ces évènements, en évaluant chacune d'elles dans l'hypothèse où l'on sait que les précédents se sont produits.*

Nous bornant au cas de deux évènements, il est clair que, si la probabilité du second ne varie pas *lorsque le premier s'est produit*, cette dernière restriction est inutile : l'on peut, dans ce cas, considérer les évènements comme simultanés.

Donnons un exemple pour mettre en évidence l'erreur que l'on peut commettre en ne tenant pas compte, dans l'évaluation de la probabilité du second évènement, du fait que le premier s'est produit.

Problème VI. *Une urne contient* 90 *boules, numérotées de* 1 *à* 90 *; on en extrait une ; quelle est la probabilité pour que son numéro soit divisible à la fois par* 6 *et par* 10 ? Il y a évidemment 1 chance sur 6 pour que le numéro soit divisible

par 6 et 1 sur 10 pour qu'il soit divisible par 10 : les probabilités respectives sont donc $\frac{1}{6}$ et $\frac{1}{10}$: ce serait une erreur de conclure que la probabilité composée est $\frac{1}{60}$; car, lorsqu'un nombre est divisible par 6, il est divisible par 2 et, par suite, pour qu'il soit divisible par 10, il suffit qu'il *soit divisible par* 5 ; la probabilité pour que le numéro tiré soit divisible par 10, *lorsque l'on sait qu'il est divisible par* 6, est donc $\frac{1}{5}$ et la vraie valeur de la probabilité composée est

$$\frac{1}{6} \times \frac{1}{5} = \frac{1}{30}.$$

Il est aisé d'arriver à ce résultat directement : pour qu'un nombre soit divisible à la fois par 6 et par 10, il faut et il suffit qu'il soit divisible par 30 ; la probabilité [1] est $\frac{1}{30}$.

14. Application aux tirages successifs d'une même urne. — Problème VII. *Une urne renferme* N *boules, parmi lesquelles* a *sont blanches. On extrait successivement* 3 *boules de l'urne. Quelle est la probabilité pour qu'elles soient blanches toutes les trois?* Nous appliquerons la règle des probabilités composées : la probabilité pour que la première boule extraite soit blanche est $\frac{a}{N}$; lorsque ce premier événement s'est produit, l'urne ne renferme plus que $N - 1$ boules parmi lesquelles $a - 1$ sont blanches ; la probabilité d'extraire une blanche au second tirage est donc $\frac{a-1}{N-1}$; après cette extraction supposée réalisée, l'urne renferme $N - 2$ boules parmi lesquelles $a - 2$ sont blanches : la probabilité de l'extraction d'une blanche au troisième tirage est donc $\frac{a-2}{N-2}$; la probabilité composée demandée

---

[1] Il n'est pas inutile d'observer que ces évaluations de probabilités sont correctes parce que 30 est divisible par 30 ; la probabilité pour qu'un nombre inférieur ou égal à $Ma + r$, r étant inférieur à a, soit divisible par a est $\frac{M}{Ma + r}$; cette probabilité se réduit à $\frac{1}{a}$ lorsque $r = 0$; elle est très voisine de $\frac{1}{a}$ lorsque M est très grand, quel que soit $r < a$.

est donc

$$\frac{a}{N}\cdot\frac{a-1}{N-1}\cdot\frac{a-2}{N-2}=\frac{a(a-1)(a-2)}{N(N-1)(N-2)}.$$

Par exemple, un jeu de 32 cartes renferme 4 rois : la probabilité pour qu'en tirant successivement 3 cartes on ait 3 rois s'obtient en faisant $N=32$, $a=4$: elle est

$$\frac{4\cdot 3\cdot 2}{30\cdot 31\cdot 32}=\frac{1}{10\cdot 31\cdot 8}=\frac{1}{2480}$$

on a une chance sur 2480.

De même un jeu de 52 cartes renferme 13 trefles : la probabilité pour obtenir 3 trèfles en tirant 3 cartes est

$$\frac{13\cdot 12\cdot 11}{52\cdot 51\cdot 50}=\frac{11}{17\cdot 50}=\frac{11}{850}=\frac{1}{77,27\ldots}.$$

On a un peu plus d'une chance sur 78.

Avec un jeu de 32 cartes, la probabilité analogue serait

$$\frac{8\cdot 7\cdot 6}{32\cdot 31\cdot 30}=\frac{7}{4\cdot 31\cdot 5}=\frac{7}{620}=\frac{1}{88,57\ldots}.$$

On a moins d'une chance sur 88 : la probabilité est plus faible.

Nous avons indiqué le détail des calculs, qui se font rapidement par la suppression des facteurs communs au numérateur et au dénominateur.

Problème VIII. *Une urne renferme N boules, parmi lesquelles a sont blanches et les autres noires : quelle est la probabilité d'amener en deux tirages une blanche et une noire ?*

On peut traiter ce problème par la combinaison des théorèmes des probabilités totales et composées : l'éventualité désirée peut se produire, soit par l'extraction d'une boule blanche au premier tirage et d'une noire au second, soit par l'extraction d'une noire au premier et d'une blanche au second : c'est donc une probabilité totale et sa valeur est

$$\frac{a}{N}\cdot\frac{N-a}{N-1}+\frac{N-a}{N}\cdot\frac{a}{N-1}=\frac{2a(N-a)}{N(N-1)}.$$

Le résultat est ici obtenu facilement, grâce à la simplicité du problème, mais dans la plupart des questions de ce genre, il est préférable de recourir à l'énumération complète des cas, par les procédés de l'analyse combinatoire. En voici un exemple.

PROBLÈME IX. *Quelle est la probabilité d'avoir d'entrée trois atouts à l'écarté?* Rappelons que l'écarté se joue avec 32 cartes, dont l'une se trouve retournée et indique l'atout : chaque joueur reçoit 5 cartes. Les 5 cartes sont donc choisies au hasard parmi les 31 cartes autres que la retournée, dont 7 sont des atouts. Le nombre de jeux possibles, lorsque l'atout est fixé, est donc égal au nombre de manières dont on peut choisir 5 cartes parmi 31 cartes, c'est-à-dire à

$$\frac{31.30.29.28.27}{1.2.3.4.5}$$

c'est le nombre des cas possibles.

Parmi ces jeux, on obtiendra le nombre de ceux qui renferment 3 atouts en multipliant le nombre des manières de choisir 3 atouts parmi les 7 atouts par le nombre de manières de choisir 2 cartes quelconques parmi les 24 cartes qui ne sont pas des atouts ; on obtient ainsi pour le nombre des cas favorables

$$\frac{7.6.5}{1.2.3}\cdot\frac{24.23}{1.2}$$

La probabilité demandée est égale au quotient du nombre de cas favorables par le nombre des cas possibles ; c'est-à-dire

$$\frac{7.6.5.24.23}{31.30.29.28.27}\cdot\frac{1.2.3.4.5}{1.2.3.1.2}$$

Des réductions immédiates donnent ($6\times 5=30$; $7\times 4=28$) :

$$\frac{24.23.5}{31.29.27.2}=\frac{4.23.5}{31.29.9}=\frac{460}{8091}=\frac{1}{17,5\ldots}$$

La probabilité demandée est comprise entre $\frac{1}{17}$ et $\frac{1}{18}$.

15. Autres problèmes.

Problème X. *Etudier le problème de la poule* (Problème III) *par la méthode des probabilités composées.*

La première partie se joue entre A et B ce que nous désignerons par la formule

(1) AB

Si A gagne, B sort et est remplacé par C ; sinon c'est A qui sort et est remplacé par C ; nous avons donc pour la seconde partie les deux formules

(2) AC′ BC′

dans lesquelles l'accent indique que le joueur C vient de rentrer. Si A gagne dans la combinaison AC′, le jeu est terminé ; sinon A sort et est remplacé par B ; on a donc pour la troisième partie les combinaisons

(3) CB′ CA′

On aurait de même pour les parties suivantes

(4) BA′ AB′
(5) AC′ BC′
(6) CB′ CA′

qui, comme on le voit se reproduisent périodiquement.

A chaque partie à partir de la deuxième, il y a une chance sur deux pour que le jeu cesse et une chance sur deux pour qu'il continue. La probabilité pour que le jeu se termine à la seconde partie est $\frac{1}{2}$; pour qu'il se termine à la troisième $\frac{1}{4}$; pour qu'il se termine à la quatrième $\frac{1}{8}$, etc.

Quelle est la probabilité de chacune des hypothèses figurées dans les formules ci-dessus ? A chaque partie, si le jeu n'est pas terminé, ces hypothèses sont toutes deux également probables, cette probabilité est donc $\frac{1}{4}$ pour la seconde partie, $\frac{1}{8}$ pour la quatrième, etc. D'autre part, la probabilité du gain d'un joueur

déterminé dans l'une de ces hypothèses est égale à 0 si ce joueur est désigné par une lettre accentuée et à $\frac{1}{2}$ s'il est désigné par une lettre non accentuée, c'est-à-dire s'il a gagné la partie précédente ; la probabilité pour qu'un joueur gagne par l'une des combinaisons où il est désigné par une lettre non accentuée est donc $\frac{1}{4}$ si la combinaison appartient à la seconde partie, $\frac{1}{8}$ à la troisième, $\frac{1}{16}$ à la quatrième, etc.

Les probabilités de gain aux diverses parties sont donc

deuxième partie $\frac{1}{4}$ pour A $\frac{1}{4}$ pour B

troisième — $\frac{1}{8} + \frac{1}{8} = \frac{1}{4}$ pour C

quatrième — $\frac{1}{16}$ pour A, $\frac{1}{16}$ pour B

cinquième — $\frac{1}{32}$ pour A, $\frac{1}{32}$ pour B

sixième — $\frac{1}{64} + \frac{1}{64} = \frac{1}{32}$ pour C

. . . . . .

$(3n+1)^{\text{ième}}$ partie $\frac{1}{2^{3n+1}}$ pour A, $\frac{1}{2^{3n+1}}$ pour B

$(3n+2)^{\text{ième}}$ — $\frac{1}{2^{3n+2}}$ pour A, $\frac{1}{2^{3n+2}}$ pour B

$(3n+3)^{\text{ième}}$ — $\frac{1}{2^{3n+3}} + \frac{1}{2^{3n+3}} = \frac{1}{2^{3n+2}}$ pour C

La probabilité pour A ou pour B de gagner au bout de $3n$ parties au plus est donc

$$\frac{1}{4} + \frac{1}{16} + \frac{1}{32} + \frac{1}{128} + \frac{1}{256} + \dots + \frac{1}{2^{3n-2}} + \frac{1}{2^{3n-1}}$$

c'est-à-dire

$$\frac{1}{4}\left(3 + \frac{1}{4} + \frac{1}{16} + \frac{1}{32} + \dots + \frac{1}{2^{3n-4}}\right)$$

ou bien, en sommant la progression géométrique

$$\frac{1}{4}+\frac{3}{2}\cdot\frac{\frac{1}{16}-\frac{1}{2^{3n+1}}}{1-\frac{1}{8}}=\frac{1}{4}+\frac{3}{28}-\frac{3}{7\cdot 2^{3n-1}}=\frac{5}{14}-\frac{3}{7\cdot 2^{3n-1}}$$

La probabilité pour C de gagner au bout de $3n$ parties au plus est

$$\frac{1}{4}+\frac{1}{32}+\cdots+\frac{1}{2^{3n-1}}=\frac{\frac{1}{4}-\frac{1}{2^{3n+2}}}{1-\frac{1}{8}}=\frac{2}{7}-\frac{1}{7\cdot 2^{3n-1}}$$

Lorsque n augmente indéfiniment, c'est-à-dire lorsqu'on ne limite pas à l'avance le nombre des parties, ces probabilités tendent vers les valeurs $\frac{5}{14}$ et $\frac{2}{7}$ que nous avons trouvées au n° **7**. On peut remarquer que la probabilité pour que l'un des trois joueurs A, B, C gagne au bout ne $3n$ parties *au plus* est

$$\frac{5}{14}-\frac{3}{7\cdot 2^{3n-1}}+\frac{5}{14}-\frac{3}{7\cdot 2^{3n-1}}+\frac{2}{7}-\frac{1}{7\cdot 2^{3n-1}}=1-\frac{1}{2^{3n-1}}$$

Elle est égale à l'unité, diminuée de la probabilité pour que le jeu ne soit pas terminé au bout de $3n$ parties.

L'étude précédente est plus longue, mais plus instructive que la méthode du n° **7** ; elle permet, par exemple, de résoudre les problèmes tels que le suivant.

PROBLÈME XI. *Quelle est l'espérance mathématique des joueurs* A *et* C *si la somme à toucher par le gagnant, au lieu d'être fixe, est égale au nombre des parties jouées.*

L'espérance mathématique de A sera

$$\frac{2}{4}+\frac{4}{16}+\frac{5}{32}+\frac{7}{128}+\cdots+\frac{3n-2}{2^{3n-2}}+\frac{3n-1}{2^{3n-1}}+\cdots$$

c'est-à-dire

$$\frac{1}{2} + \sum_{n=3}^{\infty} \frac{9n-5}{2^{3n-1}} = \frac{1}{2} + \frac{9}{4}\sum_{2}^{\infty} \frac{n}{8^n} - \frac{5}{4}\sum_{3}^{\infty} \frac{1}{8^n}$$

$$= \frac{1}{2} + \frac{9}{4}\cdot\frac{15}{49} - \frac{5}{4}\cdot\frac{1}{7}$$

$$= \frac{1}{2} + \frac{25}{49} = 1 - \frac{1}{98}$$

L'espérance mathématique de C sera

$$\frac{3}{4} + \frac{6}{32} + \ldots + \frac{3n}{2^{3n-1}} + \ldots$$

c'est-à-dire

$$\frac{3}{4}\sum_{1}^{\infty} \frac{n}{8^{n-1}} = \frac{3}{4}\cdot\frac{8^2}{7^2} = \frac{48}{49} = 1 - \frac{1}{49}$$

Elle est inférieure à celle de A, mais le désavantage est beaucoup moins grand qu'avec la première convention.

La somme des espérances mathématiques des 3 joueurs A, B, C est

$$1 + \frac{1}{98} + 1 + \frac{1}{98} + 1 - \frac{1}{49} = 3$$

On peut vérifier à priori ce résultat. En effet, la probabilité pour que la partie se termine après la $n^{ième}$ partie est $\frac{1}{2^{n-1}}$ et le gain est alors n : l'espérance mathématique totale des trois joueurs est donc (1)

$$\sum_{2}^{\infty} \frac{n}{2^{n-1}} = 3$$

Problème XII. *Quel avantage devraient réserver les deux premiers joueurs au troisième pour que le jeu de la poule soit équitable, en supposant les mises égales ?* Nous supposons que,

(1) Nous utilisons, ici et dans le calcul précédent, la formule

$$\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + 4x^3 + \ldots$$

par un procédé quelconque la probabilité de gain de C est un nombre p supérieur à $\frac{1}{2}$ lorsque C joue avec A ou avec B : lorsque A et B jouent entre eux, la probabilité de gain de chacun d'eux est $\frac{1}{2}$: on demande de déterminer p de manière que le jeu soit équitable, c'est-à-dire de manière que la probabilité de gain final soit $\frac{1}{3}$ pour chaque joueur.

Si l'on se reporte aux tableaux du problème on voit que A joue contre C les parties de rang $3n+2$ et contre B les parties de rang $3n+1$: si la probabilité du gain de A contre C est $1-p$, sa probabilité totale de gain sera

$$(1-p)\sum_{1}^{\infty}\frac{1}{2^{3n+2}}+\frac{1}{2}\sum_{2}^{\infty}\frac{1}{2^{3n+1}}$$

Nous laisserons au lecteur le soin de calculer cette expression : il vérifiera que ce calcul conduit au même résultat que le suivant, plus simple. La probabilité de gain de C est multipliée par $2p$, puisque dans chaque partie jouée par C cette probabilité est p au lieu d'être $\frac{1}{2}$; la probabilité totale de gain de C est donc égale au produit de $\frac{2}{7}$ par $2p$: elle est donc $\frac{4p}{7}$ et, comme elle doit être égale à $\frac{1}{3}$, on obtient

$$\frac{4p}{7}=\frac{1}{3}$$

c'est-à-dire

$$p=\frac{7}{12}$$

tel doit être l'avantage réservé à C : il doit avoir 7 chances sur 12 contre 5 sur 12 à A ou à B, lorsqu'il joue avec l'un de ces joueurs.

CHAPITRE III

FORMULES D'APPROXIMATION

16. But de ce chapitre. — Nous avons déjà vu que, pour avoir les valeurs numériques de diverses probabilités, on est amené à calculer les valeurs de combinaisons, valeurs dans lesquelles interviennent les produits des m premiers nombres entiers. Nous avons trouvé en effet,

$$C_m^p = \frac{m!}{p!\,(m-p)!}.$$

On exprimerait aussi avec $m!$ les expressions de la forme $1 \cdot 3 \cdot 5 \ldots (2m-1)$ c'est-à-dire le produit des nombres impairs ; on a :

$$1 \cdot 3 \cdot 5 \ldots (2m-1) = \frac{(2m)!}{2 \cdot 4 \cdot 6 \ldots 2m} = \frac{(2m)!}{2^m \cdot m!}.$$

On conçoit donc l'intérêt qui s'attache au calcul rapide de $m!$, même si ce calcul est seulement approché. Il importe de préciser ce que nous entendrons par « calcul approché ». Les « valeurs approchées » que nous indiquerons par $m!$ différeront en réalité beaucoup de la valeur exacte, en ce sens que la *différence* entre la valeur approchée et la valeur exacte sera extrêmement grande, d'autant plus grande que m sera plus grand ; mais le *rapport* de la valeur approchée à la valeur exacte différera très peu de l'unité, et en différera d'autant moins que m sera plus grand. En d'autres termes, si l'*erreur absolue* est très grande,

l'erreur relative sera très faible. Or c'est cette erreur relative qui importe, car les probabilités sont obtenues comme *quotient* de produits de factorielles.

17. Formule de Stirling. Première approximation. — La formule que l'on utilise pratiquement pour le calcul approché de $n!$ est la suivante, due à Stirling

$$n! = n^n e^{-n} \sqrt{2\pi n}\,(1 + \varepsilon_n).$$

Dans cette formule [1], n désigne un nombre entier quelconque, e la base des logarithmes népériens ($e = 2,7182818\ldots$), π le rapport de la circonférence au diamètre ($\pi = 3,14159\ldots$) et ε_n un nombre variable avec n, mais *qui tend vers zéro lorsque n augmente indéfiniment*. Dans certaines applications il peut être utile de savoir que le produit $n\varepsilon_n$ tend vers $\frac{1}{12}$ lorsque n tend vers l'infini. On posera $12n\varepsilon_n = 1 + \theta_n$.

18. Application à l'égalité des joueurs à pile ou face. — Pour en donner de suite une application, proposons-nous de calculer la probabilité d'amener n fois pile en $2n$ parties de pile ou face : d'après ce qui précède cette probabilité est

$$\frac{1}{2^{2n}} C^n_{2n} = \frac{(2n)!}{2^{2n}(n!)^2}$$

En utilisant la formule du n° **17** nous obtenons pour la valeur cherchée

$$\frac{(2n)^{2n} \cdot e^{-2n} \cdot \sqrt{4\pi n} \cdot \left(1 + \frac{1 + \theta_{2n}}{24n}\right)}{2^{2n} \cdot n^{2n} \cdot e^{-2n} \cdot 2\pi n \left(1 + \frac{1 + \theta_n}{12n}\right)^2}$$

ou, en simplifiant, et désignant $1 + \varepsilon_n$ un facteur qui tend vers 1 pour n infini (ε_n tend vers zéro) ;

$$\frac{1 + \varepsilon_n}{\sqrt{\pi n}}.$$

---

[1] Nous renverrons pour la démonstration de cette formule, aux ouvrages d'analyse ; cette démonstration n'est pas utile à connaître pour l'emploi que nous avons à faire de la formule.

On voit que cette probabilité décroît lorsque n augmente, mais décroît très lentement.

Pour mesurer l'approximation pour de faibles valeurs de n, on peut avoir recours au triangle arithmétique ; pour $n = 5$, on a la relation

$$\frac{252}{1024} = \frac{1 + \varepsilon_5}{\sqrt{5\pi}}$$

qui donne environ

$$1 + \varepsilon_5 = 0{,}99$$

c'est-à-dire que l'erreur *relative* ε_5 est négative et égale à environ $\frac{1}{100}$.

Si l'on remarque que $\sqrt{\pi}$ est égal à environ $1{,}772$ on voit que pour $n = 100$ la probabilité d'amener 100 fois pile sur 200 parties est

$$\frac{1}{10 \times 1{,}772} = \frac{1}{17{,}72}$$

c'est-à-dire à environ $\frac{1}{18}$. La probabilité d'amener 10 000 fois pile sur 20 000 parties est de même sensiblement égale à $\frac{1}{177}$.

On voit que la probabilité pour un joueur à pile ou face, de se trouver sans gain ni perte, va en diminuant lorsque le nombre de coups *fixé à l'avance* va en augmentant ; la probabilité pour que la partie s'équilibre au 200ième coup est $\frac{1}{18}$; la probabilité pour qu'elle s'équilibre au 20 000ème est seulement $\frac{1}{177}$; la probabilité pour qu'elle s'équilibre au 2 000 000ième est seulement $\frac{1}{1772}$. On commettrait une erreur grave, analogue à celle que nous avons signalée au n° 9, si l'on disait : comme il arrive assez souvent que la partie s'équilibre, on peut toujours supposer que le jeu recommence à ce moment là ; par conséquent la probabilité reste constamment la même. En réalité, les longues séries de coups avantageuses dans leur ensemble à l'un des joueurs sont assez peu probables ; mais, lorsque le jeu dure assez longtemps,

elles finissent tout de même par se produire, et elles peuvent alors empêcher l'équilibre de se produire pendant longtemps.

On se rend nettement compte de cette diminution des chances d'équilibre en calculant l'espérance mathématique d'un spectateur du jeu auquel les deux joueurs verseraient une certaine somme fixe chaque fois que l'équilibre se produirait. Nous allons faire ce calcul en supposant exacte la formule approchée : l'erreur totale est très faible, car l'erreur n'est sensible que pour les petites valeurs de n.

L'espérance mathématique est la somme des espérances mathématiques dans chacune des hypothèses où l'égalité peut se produire, c'est-à-dire apres 2 parties, 4 parties, 6 parties, ..., $2n$ parties. Sa valeur est donc, pour les $2n$ premières parties

$$\frac{1}{\sqrt{\pi}}\left(1+\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}+\ldots\ldots+\frac{1}{\sqrt{n}}\right).$$

On obtient une valeur très approchée de la quantité entre parenthèses en la remplaçant par l'intégrale

$$\int_0^n \frac{dx}{\sqrt{x}} = 2\sqrt{n}$$

et l'on obtient ainsi, pour valeur approchée de l'espérance mathématique cherchée

$$\frac{2}{\sqrt{\pi}}\sqrt{n}$$

c'est-à-dire à peu près $1,128\sqrt{n}$. Ainsi pour 200 parties l'espérance mathématique est 11 ; pour 800 parties, elle est seulement le double, c'est-à-dire 22 ; pour deux millions de parties, elle est 1128 environ.

Supposons que l'on joue devant Paul des séries de 1000 parties, chaque série étant indépendante des précédentes, c'est-à-dire que, la série terminée, on suppose que l'on commencera à noter les résultats à zéro. Si Paul verse 20 francs pour chaque série, et doit toucher 1 franc chaque fois que l'égalité se produit, il fera un marché très vraisemblablement avantageux on peut

même dire sûrement avantageux, en vertu de la loi des grands nombres (voir le chapitre suivant) si le nombre des séries est considérable : si ce nombre de séries est 2 000, ce qui porte à deux millions le nombre total des parties jouées, il aura versé 50 000 francs et réalisé un bénéfice de plusieurs milliers de francs. Mais si les deux millions de parties étaient regardées comme formant une série unique, la promesse faite à Paul de lui verser 1 franc chaque fois que l'égalité se produit ne vaudrait pas plus de 1 128 francs : en la payant 2 000 francs, il ferait un très mauvais marché et ce serait une folie de la payer 50 000 francs. On voit nettement sur cet exemple concret combien serait grossière l'erreur qui consisterait, sous prétexte que l'on revient à l'égalité en moyenne plus de 1 000 fois sur deux millions de parties, à penser que l'on peut sans inconvénient, décomposer cette série en séries partielles comprenant chacune 2 000 parties environ : en réalité, l'irrégularité des retours à l'égalité est beaucoup plus grande ; il arrivera fréquemment ceci : pendant quelques milliers de parties, il s'en produira une vingtaine par milliers de parties et Paul s'imaginera qu'il va s'enrichir rapidement ; puis il se produira un écart exceptionnel, l'un des joueurs ayant, comme l'on dit vulgairement, « de la chance » pendant la durée de quelques centaines de parties et, cet écart exceptionnel s'étant une fois produit, le retour à l'égalité deviendra peu vraisemblable, puisqu'il faudra pour le réaliser un autre écart exceptionnel de sens contraire ; il pourra alors arriver que, pendant des milliers de parties, Paul ne touche pas un centime : l'égalité finira cependant probablement par se produire à nouveau et sera suivie vraisemblablement d'une période de gains pour Paul, et ainsi de suite. Ajoutons que, si nous appelons, pour fixer les idées, écarts exceptionnels les écarts qui dépassent 50, il arrivera, si l'on joue des centaines de milliers de parties, que parmi ces écarts il y en aura qui seront encore exceptionnels, qui dépasseront 200 ou 300 ; et ils seront vraisemblablement suivis d'une période extrêmement longue, pouvant atteindre cent mille parties, pendant laquelle l'égalité ne se produira pas.

19. Deuxième approximation. — On peut, au lieu de la formule du n° **17**, obtenir une approximation plus précise : si

l'on pose

$$n! = n^n e^{-n}\sqrt{2\pi n}\, f(n) \tag{1}$$

on démontre les relations suivantes

$$\log f(n+1) - \log f(n) = -\frac{1}{12n^2} + \frac{1}{12n^3} - \frac{3}{10n^4} + \cdots + \frac{(-1)^{p-1}(p-1)}{2p(p-1)n^p} + \cdots \tag{2}$$

$$\lim_{n=\infty} f(n) = 1. \tag{3}$$

Ces trois relations permettent d'obtenir une approximation aussi grande que l'on veut : la formule du n° **17** s'obtient en réduisant le second membre de la formule (2) à son premier terme ; pour tenir compte des termes suivants, nous poserons en utilisant le résultat déjà rappelé

$$f(n) = 1 + \frac{1}{12n} + \frac{1-\theta_n}{2n^2}, \tag{4}$$

θ_n tendant vers zéro lorsque n augmente indéfiniment et α étant un coefficient à calculer ; on a

$$\log f(n) = \frac{1}{12n} + \frac{1-\theta_n}{2n^2} - \frac{1}{2}\left(\frac{1}{12n} + \frac{1+\theta_n}{2n^2}\right)^2 + \cdots$$
$$= \frac{1}{12n} + \left(\frac{1}{2} - \frac{1}{288}\right)\frac{1}{n^2} + \cdots \tag{5}$$

les termes non écrits étant en $\frac{1}{n^3}$. On a donc

$$\log f(n+1) - \log f(n) = \frac{1}{12}\left(\frac{1}{n+1} - \frac{1}{n}\right) + \left(\frac{1}{2} - \frac{1}{288}\right)\left(\frac{1}{(n+1)^2} - \frac{1}{n^2}\right) + \cdots$$

Or

$$\frac{1}{n+1} - \frac{1}{n} = -\frac{1}{n^2} + \frac{1}{n^3} - \frac{1}{n^4} + \cdots$$

$$\frac{1}{(n+1)^2} - \frac{1}{n^2} = -\frac{2}{n^3} + \frac{3}{n^4} - \frac{4}{n^5} + \cdots$$

Par suite, en n'écrivant que les termes en $\frac{1}{n^2}$ et $\frac{1}{n^3}$

$$\log f(n-1) - \log f(n) = -\frac{1}{12n^2} + \left(\frac{1}{12} - \frac{2}{\alpha} + \frac{1}{144}\right)\frac{1}{n^3} + \dots$$

et l'identification avec la formule (2) donne la valeur de α :

$$\alpha = 288.$$

Nous ne recrirons pas ici la formule (4) car nous allons résumer les formules essentielles obtenues dans ce chapitre.

20. Résumé des formules. — On peut, dans certains cas, se contenter de prendre

$$(1) \qquad n! = n^n e^{-n}\sqrt{2\pi n}\,(1+\varepsilon_n)$$

et l'on sait que ε_n tend vers zéro lorsque n augmente indéfiniment ; cette formule s'écrit aussi

$$(2) \qquad \log(n!) = \left(n+\frac{1}{2}\right)\log n - n + \log\sqrt{2\pi} + \varepsilon'_n$$

ε'_n tendant vers zéro lorsque n augmente indéfiniment.

On a une première approximation de ε_n en prenant

$$\varepsilon_n = \frac{1+\theta_n}{12n}$$

Ce qui donne les formules

$$(3) \qquad n! = n^n e^{-n}\sqrt{2\pi n}\left(1+\frac{1+\theta_n}{12n}\right)$$

$$(4) \qquad \log(n!) = \left(n+\frac{1}{2}\right)\log n - n + \log\sqrt{2\pi} - \frac{1+\theta'_n}{12n}$$

Enfin une seconde approximation donne

$$(5) \qquad n! = n^n e^{-n}\sqrt{2\pi n}\left(1+\frac{1}{12n}+\frac{1+\gamma_n}{288n^2}\right).$$

Il est remarquable que la formule (4) donne une approximation aussi grande que la formule (5), si l'on y suppose que $n\theta'_n$ tend vers zéro lorsque n augmente indéfiniment : cela résulte de la

formule 5 du paragraphe précédent ou l'on fait $\alpha = 288$; le terme en $\frac{1}{n^2}$ disparait. On obtient, par un calcul que nous omettons une *troisième* approximation pour $\log n!$

$$(6)\quad \log(n!) = \left(n + \frac{1}{2}\right)\log n - n + \log\sqrt{2\pi} + \frac{1}{12n} - \frac{1 + \theta_n}{360n^3}$$

Signalons enfin la formule suivante, due à M. Forsyth

$$(7)\qquad n! = \sqrt{2\pi}\left[\frac{\sqrt{n^2 + n + \frac{1}{6}}}{e}\right]^{n+\frac{1}{2}}$$

qui équivaut à la formule 6 dans laquelle $\frac{1 + \theta_n}{360n^3}$ serait remplacée par $\frac{1}{1440n^3}$; l'erreur comprise sur $\log(n!)$ par l'emploi de la formule (7) est donc à peu près égale à $\frac{1}{2520n^3}$, tandis que l'erreur commise par l'emploi de la formule (4) est environ $\frac{1}{360n^3}$.

Pratiquement, les formules (1) et (2) suffisent pour les applications numériques courantes : les formules (3), (4), (5), (6) sont utilisées dans les recherches théoriques.

CHAPITRE IV

ÉTUDE APPROFONDIE DU JEU DE PILE OU FACE

21. Formule fondamentale. — Cherchons qu'elle est la probabilité pour que, en $2m$ parties de pile ou face, on amène pile $m - h$ fois et face $m + h$ fois : cette probabilité p est donnée par la formule

$$p = \frac{1}{2^{2m}} \frac{(2m)!}{(m-h)!\,(m+h)!}.$$

Nous allons calculer $\log p$ en évaluant les factorielles au moyen de la formule (1) nous aurons, en désignant par θ un nombre compris entre 0 et 1, nombre qui n'est pas le même dans les diverses formules :

$$\log p = -2m\log 2 + \left(2m + \frac{1}{2}\right)\log(2m) - 2m + \log\sqrt{2\pi} + \frac{1+\theta}{24m}$$
$$- \left(m - h + \frac{1}{2}\right)\log(m-h) + (m-h) - \log\sqrt{2\pi} - \frac{1+\theta'}{12(m-h)}$$
$$- \left(m + h + \frac{1}{2}\right)\log(m+h) + (m+h) - \log\sqrt{2\pi} - \frac{1+\theta'}{12m+h}$$

Pour simplifier cette formule, on remarque que l'on a

$$\log(2m) = \log 2 + \log m$$

et que le coefficient $2m + \frac{1}{2}$ de $\log m$ peut être décomposé

comme il suit

$$2m + \frac{1}{2} = -\frac{1}{2} + (m - h + \frac{1}{2}) + \left(m + h + \frac{1}{2}\right).$$

On peut des lors écrire

$$\log p = -\frac{1}{2}\log m - \log\sqrt{\pi} + \varepsilon_m$$
$$+ \left(m - h + \frac{1}{2}\right)\left[\log m - \log(m - h)\right]$$
$$+ \left(m + h + \frac{1}{2}\right)\left[\log m - \log(m + h)\right]$$

en désignant par ε_m la somme

$$\varepsilon_m = \frac{1 + \theta}{2(m} - \frac{1 + \theta}{12(m - h} - \frac{1 + \theta}{12(m + h)}$$

qui est très petite lorsque m et $m - h$ sont très grands.

Or on a

$$\log m - \log(m - h) = -\log\left(1 - \frac{h}{m}\right) = \frac{h}{m} + \frac{h^2}{2m^2} + \frac{h^3}{3m^3} +$$
$$\log m - \log(m + h) = -\log\left(1 + \frac{h}{m}\right) = -\frac{h}{m} + \frac{h^2}{2m^2} - \frac{h^3}{3m^3} + \ldots$$

Ces séries sont convergentes du moment que l'on suppose $h < m$. On en conclut

$$\left(m - h + \frac{1}{2}\right)\left[\log m - \log(m - h)\right] = h + \frac{h^2}{2m} + \frac{h^3}{3m^2} + \ldots$$
$$- \frac{h^2}{m} - \frac{h^3}{2m^2} - \ldots$$
$$+ \frac{1}{2}\frac{h}{m} + \frac{1}{2}\frac{h^2}{2m^2} + \ldots$$

Le produit $\left(m + h + \frac{1}{2}\right)\left[\log m - \log(m + h)\right]$ se déduit du précédent par le changement de h en $-h$: on en conclut finalement

$$\log p = -\frac{1}{2}\log m - \log\sqrt{\pi} + \varepsilon_m - \frac{h^2}{m} + \frac{h^2}{2m^2} - \ldots$$

Les termes non écrits renferment m en dénominateur avec un exposant au moins égal à 3 et l'exposant de h, qui peut y figurer en numérateur, dépasse au plus d'une unité l'exposant de m.

Comme h est toujours inférieur à m ces termes sont inférieurs en valeur absolue au terme négatif $-\frac{h^2}{m}$: d'ailleurs lorsque h^2 est grand par rapport à m, ce terme négatif est grand en valeur absolue et par suite $\log p$ devient extrêmement petit : c'est ce qui se produit si l'on suppose, pour fixer les idées $h^3 > m^2$: on a alors

$$\frac{h^2}{m} > \frac{h^2}{h\sqrt{h}} > \sqrt{h}$$

et $p < e^{-\sqrt{h}} < e^{-\sqrt{m}}$.

Si, par exemple, $m = 1\,000$, et $h > 100$, p est inférieur à e^{-10} si $m = 1\,000\,000$ et $h > 10\,000$, p est inférieur à e^{-100}. La valeur absolue de p est ici assez faible pour que l'on n'ait pas besoin d'une grande précision relative dans son calcul ; aussi allons-nous continuer le calcul de la formule d'approximation en supposant $h < m^2$: elle devient alors, en négligeant tous les termes qui renferment m^2 en dénominateur

$$\log p = -\frac{1}{2}\log m\pi + \varepsilon_m - \frac{h^2}{m}$$

et par suite

$$p = \frac{1}{\sqrt{m\pi}} e^{-\frac{h^2}{m}}(1 + \varepsilon_m),$$

en désignant par ε_m comme par ε_m un nombre qui tend vers zéro lorsque m augmente indéfiniment.

Cette formule généralise celle que nous avions obtenue au n° 18 en supposant $h = 0$; elle est d'une très grande importance : aussi n'est-il pas inutile de bien préciser qu'elle est une conséquence analytique immédiate des formules des combinaisons, et qu'elle ne renferme aucune autre hypothèse.

22. Unité d'écart. — Il est commode de transformer la formule en y posant

$$h = \lambda \sqrt{m}$$

le nombre h s'appelle *l'écart*.

La quantité $\sqrt{m}$ prend le nom *d'unité d'écart* et λ s'appelle *l'écart relatif* : on obtient alors

$$p = \frac{1}{\sqrt{m\pi}} e^{-\lambda^2} (1 + \varepsilon_m)$$

et l'on voit que le terme exponentiel ne dépend que de la valeur de *l'écart relatif* λ, ce qui justifie son introduction.

Lorsque h et m sont des nombres très grands, on peut faire varier h de plusieurs unités sans que λ varie sensiblement. Désignons par h_1 et h_2 deux nombres entiers voisins et cherchons la probabilité pour que l'écart h soit compris entre h_1 et h_2 ; d'une manière plus précise pour que l'on ait

$$h_1 \leq h < h_2$$

Nous poserons

$$\lambda_1 = h_1 \sqrt{m} \qquad \lambda_2 = h_2 \sqrt{m}$$

et nous supposerons λ_1 et λ_2 assez voisins pour que l'on puisse, sans erreur sensible, les remplacer par un même nombre λ dans l'exponentielle $e^{-\lambda^2}$ $\lambda_1 < \lambda < \lambda_2$: la probabilité cherchée est alors égale à la somme d'autant de termes qu'il y a de valeurs de h ; comme ces termes sont égaux sa valeur totale est

$$P = (h_2 - h_1) \frac{1}{\sqrt{m\pi}} e^{-\lambda^2} (1 + \varepsilon_m)$$

ou, en remplaçant h_2 et h_1 par leurs valeurs et remarquant que le facteur $\sqrt{m}$ disparait

$$P = \frac{\lambda_2 - \lambda_1}{\sqrt{\pi}} e^{-\lambda^2} (1 + \varepsilon_m).$$

Cette formule a ceci de très remarquable qu'elle renferme seulement les écarts relatifs λ. Elle fait connaître une homogénéité spéciale aux questions de probabilité.

Prenons par exemple $m = 200$, $h_1 = 5$, $h_2 = 7$: on se propose donc de rechercher quelle est la probabilité pour que le nombre de parties amenant pile, sur 200 parties de pile ou face, soit 105 ou 106, c'est-à-dire soit supérieur ou égal à 105 et inférieur à 107 : on a ici

$$\sqrt{m} = 10 \qquad \lambda_1 = 0{,}5 \qquad \lambda_2 = 0{,}7.$$

Les valeurs de λ qui interviennent dans l'exponentielle sont 0,5 et 0,6 dont les carrés sont 0,25 et 0,36 : on peut prendre environ 0,3 ce qui donne

$$\frac{e^{-\lambda^2}}{\sqrt{\pi}} = \frac{e^{-0{,}3}}{\sqrt{\pi}} = 0{,}42.$$

La probabilité demandée est donc à l'approximation des chiffres écrits

$$P = (\lambda_2 - \lambda_1) \times 0{,}42 = 0{,}2 \times 0{,}42 = 0{,}084.$$

Prenons maintenant $2m = 20\,000$, $h_1 = 50$, $h_2 = 70$: les valeurs de λ_1 et de λ_2 seront les mêmes, et par suite *la probabilité sera la même*. La probabilité ici définie est la probabilité, en 20 000 parties de pile ou face, d'amener un nombre de fois pile égal à [1] à 10050, ou 10051, ou 10052.... ou 10069, c'est-à-dire supérieur ou égal à 10050 et inférieur à 10070.

Cet exemple précise bien les conditions d'homogénéité ; il ne serait pas exact de dire que la probabilité d'amener 10050 fois pile en 20 000 parties est la même que celle d'amener 105 fois pile en 200 parties : elle est en réalité environ 10 fois plus faible ; à la probabilité d'amener 105 fois pile en 200 parties correspond en réalité la probabilité d'amener en 20 000 parties un nombre de fois pile compris entre 10050 et 10059 (ou, si l'on préfère, entre 10045 et 10054), c'est-à-dire la somme de 10 probabilités à peu près égales entre elles.

---

[1] Il aurait été plus correct de prendre $h_1 = 45$, $h_2 = 65$, dont la moyenne est 55 de même que la moyenne de $h_1 = 5$ et $h_2 = 6$ est 5,5. Mais nous n'insistons pas sur ce point, car on emploie pratiquement presque toujours les intégrales comme nous l'expliquerons tout à l'heure.

Le procédé de calcul que nous venons d'indiquer ne fournit qu'une approximation assez médiocre, mais qui cependant est souvent suffisante ; pour en faciliter l'emploi nous donnons ci-dessous une table des valeurs à un millième près du facteur $\frac{1}{\sqrt{\pi}} e^{-\lambda^2}$ pour certaines valeurs de λ.

λ	$\frac{1}{\sqrt{\pi}} e^{-\lambda^2}$	λ	$\frac{1}{\sqrt{\pi}} e^{-\lambda^2}$
0,01	0,564	0,4	0,481
0,02	0,564	0,5	0,439
0,03	0,564	0,6	0,394
0,04	0,563	0,7	0,346
0,05	0,563	0,8	0,297
0,06	0,562	0,9	0,251
0,07	0,561	1	0,208
0,08	0,561	1,5	0,059
0,09	0,560	2	0,010
0,1	0,559	2,5	0,001
0,2	0,542		
0,3	0,516		

23. Emploi de la variable continue. — Nous avons vu que la probabilité pour que l'écart *relatif* λ soit compris entre λ_1 et λ_2 est donnée, à une approximation suffisante, par la formule

$$(1) \qquad P = (\lambda_2 - \lambda_1) \frac{1}{\sqrt{\pi}} e^{-\lambda^2}.$$

Dans cette formule les écarts relatifs λ_1 et λ_2 sont donnés par les formules

$$h_2 = \lambda_2 \sqrt{m} \qquad h_1 = \lambda_1 \sqrt{m}$$

dans lesquelles h_2 et h_1 sont nécessairement des nombres entiers ; les nombres λ_1 et λ_2 ne peuvent donc pas prendre toutes les valeurs possibles.

Mais, lorsque $\sqrt{m}$ est très grand, les valeurs de λ qui varient par échelons égaux à $\frac{1}{\sqrt{m}}$, peuvent être considérées comme for-

mant une suite continue ; on est ainsi conduit à remplacer la formule (1) qui donne P par la suivante

$$P = \frac{1}{\sqrt{\pi}} \int_{\lambda_1}^{\lambda_2} e^{-\lambda^2} d\lambda. \quad (2)$$

On remarquera d'ailleurs que, en vertu du théorème de la moyenne, la valeur de cette intégrale est donnée par la formule (1) dans laquelle λ est un nombre compris entre λ_1 et λ_2. L'avantage de la formule (2), c'est qu'on n'est plus obligé d'y supposer que λ_1 et λ_2 sont très voisins l'une de l'autre, car si l'on suppose $\lambda_1 < \lambda_2 < \lambda_3$ et que l'on pose :

$$P_{1,2} = \frac{1}{\sqrt{\pi}} \int_{\lambda_1}^{\lambda_2} e^{-\lambda^2} d\lambda ;$$

$$P_{2,3} = \frac{1}{\sqrt{\pi}} \int_{\lambda_2}^{\lambda_3} e^{-\lambda^2} d\lambda.$$

on a

$$P_{1,2} + P_{2,3} = \frac{1}{\sqrt{\pi}} \int_{\lambda_1}^{\lambda_3} e^{-\lambda^2} d\lambda.$$

Ceci se généraliserait pour un nombre quelconque d'intervalles successifs de variation de λ.

La probabilité P pour que λ soit compris entre $-a$ et $+a$ est

$$P_a = \frac{1}{\sqrt{\pi}} \int_{-a}^{+a} e^{-\lambda^2} d\lambda.$$

Lorsque a augmente indéfiniment cette probabilité tend vers la limite

$$P_\infty = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\lambda^2} d\lambda.$$

Or, on démontre aisément que la valeur de cette intégrale est l'unité [1]. Ce résultat appelle quelques observations.

---

(1) Voici la démonstration la plus simple ; posons

$$J = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x^2} dx$$

Si, au lieu d'employer des formules approchées, nous avions employé des formules exactes, il est bien clair que la probabilité totale pour que l'écart relatif λ prenne toutes les valeurs qu'il est susceptible de prendre serait égale à l'unité. Or, on a posé

$$\lambda = \frac{h}{\sqrt{m}}$$

et h peut varier de $-m$ à $+m$ puisque le nombre $m+h$ des parties gagnées peut varier de 0 à $2m$. Donc λ peut varier de $-\sqrt{m}$ à $+\sqrt{m}$ et par suite λ^2 de 0 à m. Mais dès que m est un peu grand, $e^{-\lambda^2}$ devient négligeable lorsque λ^2 dépasse m : par exemple si $m = 100$, e^{-m} est inférieur à 10^{-43} ; c'est une quantité absolument négligeable dans tous les calculs. On ne commet donc *aucune erreur* en étendant l'intégration à l'intervalle $-\infty, +\infty$, au lieu de se borner à l'intervalle $-m, +m$ et il est remarquable que la valeur totale de la probabilité ainsi obtenue soit précisément égale à l'*unité*, bien que la formule utilisée soit *simplement approchée* : cela prouve que les petites erreurs entraînées par les approximations faites au n° **21** se compensent mutuellement. Cette circonstance est avantageuse pour l'application des formules, comme nous allons le voir.

---

on a aussi

$$J = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-y^2} dy ;$$

et, par suite

$$J^2 = \frac{1}{\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-x^2-y^2} dx dy$$

si l'on transforme en coordonnées polaires cette intégrale double étendue à tout le plan, il vient

$$J^2 = \frac{1}{\pi} \int_0^{\infty} \int_0^{2\pi} e^{-\rho^2} \rho \, d\rho \, d\omega = \frac{1}{\pi} \int_0^{2\pi} \frac{d\omega}{2} = 1$$

donc

$$J = 1.$$

24. Définition de la fonction $\theta(\lambda)$. — On désigne habituellement par $\theta(\lambda)$ la probabilité pour que l'écart relatif soit compris entre $-\lambda$ et $+\lambda$; on pose donc

$$\theta(\lambda) = \frac{1}{\sqrt{\pi}} \int_{-\lambda}^{+\lambda} e^{-\lambda^2} d\lambda$$

ou, ce qui revient au même

$$\theta(\lambda) = \frac{2}{\sqrt{\pi}} \int_{0}^{\lambda} e^{-\lambda^2} d\lambda.$$

On a calculé des tables de la fonction $\theta(\lambda)$, d'un usage courant dans les applications du calcul des probabilités : donnons seulement ici quelques unes des valeurs de cette fonction. Nous donnons aussi les valeurs de la différence $1 - \theta(\lambda)$; on a :

$$1 - \theta(\lambda) = \frac{2}{\sqrt{\pi}} \int_{\lambda}^{\infty} e^{-\lambda^2} d\lambda$$

λ	$\theta(\lambda)$	$1 - \theta(\lambda)$
0,1	0,1125	0,8875
0,2	0,2227	0,7773
0,3	0,3286	0,6714
0,4	0,4284	0,5716
0,4769	0,5	0,5
0,5	0,5205	0,4795
0,6	0,6039	0,3961
0,7	0,6778	0,3222
0,8	0,7421	0,2579
0,9	0,7969	0,2031
1	0,8427	0,1573
1,1	0,8802	0,1198
1,163	0,9	0,1
1,5	0,9661	0,0339
1,821	0,99	0,01
2	0,9953	0,0047
2,327	0,999	0,001
2,751	0,9999	0,0001
3	0,999979	0,000021
3,123	0,99999	0,00001
3,49	0,999999	0,000001
3,793	0,9999999	0,0000001
4	0,99999985	0,00000015

On voit avec quelle rapidité $\theta(\lambda)$ s'approche de l'unité lorsque λ augmente : pour $\lambda = 4$ la valeur $1 - \theta(\lambda)$ est égale à $\frac{1,5}{10^5}$: c'est une quantité presque toujours négligeable dans les calculs pratiques.

La table précédente permet de résoudre avec une approximation suffisante les diverses questions que l'on peut se poser sur le jeu de pile ou face. Donnons-en quelques exemples.

25. Problèmes où l'on utilise $\theta(\lambda)$.

Problème XVI. — *On joue* 200 *parties de pile ou face ; quelle est la probabilité d'amener pile plus de* 110 *fois?* On a ici, $h = 10$, $m = 100$; donc [1]

$$\lambda = \frac{h}{\sqrt{m}} = 1.$$

La probabilité pour que l'écart relatif dépasse 1 est $1 - \theta(1) = 0,157$ d'après la table ; mais on doit prendre garde que nous désirons ici que l'écart relatif dépasse 1 *dans un sens déterminé*, c'est-à-dire que nous tenons compte seulement des cas où l'on amène *pile* plus de 110 fois, et non pas de tous ceux où l'on amène pile *ou* face plus de 110 fois ; on doit donc diviser par 2 la probabilité précédente, ce qui donne 0,078. Telle est la probabilité demandée ; elle est égale environ à *un treizième*.

Problème XVII. — *Quelle est la probabilité pour que, sur* 20000 *parties de pile ou face, l'écart absolu dépasse* 300, *c'est-à-dire pour que pile soit amené moins de* 9700 *fois ou plus de* 10300 ?

L'écart relatif est ici $\frac{300}{\sqrt{10000}} = 3$: la probabilité qu'il soit dépassé est $1 - \theta(3)$, c'est-à-dire 0,00002 : il y a donc environ

(1) Il serait plus rigoureux de prendre $h = 10,5$ afin de donner à λ une valeur moyenne entre la valeur qui correspond à $h = 10$ et la valeur qui correspond à $h = 11$, nous négligerons cette correction, assez peu importante ; mais il était utile de la signaler au lecteur désireux d'effectuer les calculs avec précision.

1 chance sur 50 000 pour que se produise l'éventualité visée dans l'énoncé.

26 Ecart le plus probable, écart probable, écart médian. — La probabilité pour que l'écart soit compris entre λ_1 et λ_2 est

$$\frac{1}{2}\left[\Theta(\lambda_2) - \Theta(\lambda_1)\right] = \frac{1}{\sqrt{\pi}} \int_{\lambda_1}^{\lambda_2} e^{-\lambda^2} d\lambda$$

Lorsque l'intervalle $\lambda_2 - \lambda_1$ est constant, cette probabilité diminue lorsque la valeur absolue de λ augmente : elle est la plus grande possible pour $\lambda = 0$; c'est-à-dire que *l'écart le plus probable est l'écart nul*. Il importe de bien préciser le sens de cette constatation. Supposons que nous fassions un grand nombre de séries de 200 parties de pile ou face et que nous notions, à chaque série, le nombre de fois où sort pile. Lorsque pile sortira 100 fois, nous dirons que l'écart est nul ; lorsque pile sortira 101 fois qu'il est égal à + 1 et lorsque pile sortira 99 fois, qu'il est égal à — 1. Sur un grand nombre de séries, ce seront les écarts égaux à zéro qui seront les plus nombreux : ils sont plus nombreux que les écarts égaux à + 1, plus nombreux aussi que les écarts égaux à — 1, *mais moins nombreux que ces deux catégories réunies*. D'ailleurs, les formules des combinaisons donnent, d'une manière plus précise que les formules approchées, le rapport entre la probabilité de l'écart 0, de l'écart + 1 ; ce rapport est

$$C_{200}^{100} : C_{200}^{101} = \frac{101}{100} = 1{,}01.$$

On appelle *écart probable* ou *valeur moyenne de l'écart* la somme des produits obtenus en multipliant chacun des écarts possibles par sa probabilité. En d'autres termes, *c'est l'espérance mathématique d'une personne qui devrait recevoir une somme égale à l'écart*. Dans le calcul de l'écart probable, il importe de préciser si l'on prend la valeur algébrique de l'écart ou sa valeur absolue ; dans le premier cas, la quantité cherchée est évidemment zéro, car les écarts négatifs sont aussi probables que les écarts positifs ; aussi fait-on le plus souvent la seconde convention, c'est-à-dire que *l'écart est pris en valeur absolue*. La

valeur de l'écart probable est alors évidemment donnée par l'intégrale

$$\frac{2}{\sqrt{\pi}} \int_0^{\infty} \lambda e^{-\lambda^2} d\lambda = \frac{1}{\sqrt{\pi}}$$

car la probabilité d'un écart compris en valeur absolue entre λ_1 et λ_2 est $\frac{2}{\sqrt{\pi}} \int_{\lambda_1}^{\lambda_2} e^{-\lambda^2} d\lambda$. Ainsi *la valeur moyenne de l'écart relatif est* $\frac{1}{\sqrt{\pi}} = 0{,}5663\ldots$ La valeur de l'écart absolu s'obtient en multipliant cette valeur par l'unité d'écart $\sqrt{m}$. Ainsi sur une série de 20 000 parties, la valeur moyenne de l'écart est

$$\frac{100}{\sqrt{\pi}} = 56{,}63.$$

Telle est l'espérance mathématique d'une personne qui devrait recevoir une somme égale à l'écart : mais ce n'est pas l'écart qui se réalise le plus souvent ; l'écart qui se réalise le plus souvent est l'écart zéro.

Nous donnerons le nom d'*écart médian* à une valeur telle que l'on a des chances égales à parier qu'elle sera dépassée par l'écart ou à parier qu'elle ne sera pas dépassée. La valeur de l'écart médian correspond donc à la valeur de λ pour laquelle on a

$$\theta(\lambda) = \frac{1}{2},$$

On voit dans la table du n° **24** que cette valeur est $\lambda = 0{,}4769$. Telle est la valeur de l'écart relatif médian [1] : l'écart absolu médian est égal au produit de ce nombre par $\sqrt{m}$; en reprenant l'exemple précédent ($2m = 20\,000$), on trouve

$$47{,}69.$$

Ainsi sur 20 000 parties de pile ou face, on peut, avec chances égales, parier que l'écart dépassera 47 ou qu'il sera au plus égal

(1) On appelle quelquefois cet écart : écart moyen, mais cette appellation me paraît défectueuse, car il n'entre dans sa définition aucune idée de moyenne.

à $\frac{1}{4}$. On voit que l'écart médian est inférieur à l'écart probable ou moyen : leur rapport est égal à 0.845 environ.

27. Valeur moyenne d'une fonction quelconque. — D'une manière générale, on appelle valeur probable ou valeur moyenne d'une certaine fonction $f(\lambda)$ de l'écart λ la somme des produits des diverses valeurs possibles de cette fonction par leurs probabilités respectives. Nous désignerons cette valeur par $\mathfrak{M}|f(\lambda)|$; on a :

$$\mathfrak{M}|f(\lambda)| = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} f(\lambda)\, e^{-\lambda^2} d\lambda.$$

D'après cette définition on a :

$$\mathfrak{M}(1) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\lambda^2} d\lambda = 1$$

$$\mathfrak{M}(\lambda) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \lambda e^{-\lambda^2} d\lambda = 0$$

$$\mathfrak{M}(|\lambda|) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} |\lambda|\, e^{-\lambda^2} d\lambda = \frac{1}{\sqrt{\pi}}.$$

La première de ces relations exprime un fait évident : si une quantité est constante, sa valeur moyenne est égale à elle-même ; les deux autres équations sont l'expression de résultats déjà obtenus au paragraphe précédent.

Calculons la valeur moyenne de λ^2, c'est-à-dire du carré de l'écart ; on a

$$\mathfrak{M}(\lambda^2) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \lambda^2 e^{-\lambda^2} d\lambda.$$

Or on a [1]

$$J_2 = \int_0^{+\infty} \lambda^2 e^{-\lambda^2} d\lambda = \frac{1}{2}\sqrt{\pi}$$

---

[1] Si l'on pose, d'une manière générale

$$J_n = \int_0^{+\infty} \lambda^n e^{-\lambda^2} d\lambda$$

et l'on en conclut :

$$\mathfrak{M}(\lambda^2) = \frac{1}{2}.$$

Si au lieu des écarts relatifs on considérait les écarts absolus $h = \lambda\sqrt{m}$, on trouverait

$$\mathfrak{M}(h^2) = \frac{m}{2}.$$

Il est intéressant de comparer la moyenne des carrés des écarts avec le carré de leurs valeurs absolues ; on a :

$$\frac{[\mathfrak{M}(|h|)]^2}{\mathfrak{M}(h^2)} = \frac{\left(\frac{\sqrt{m}}{\sqrt{\pi}}\right)^2}{\frac{m}{2}} = \frac{2}{\pi}.$$

Ce résultat remarquable, qui ne renferme que les rapports des écarts, peut être soumis à une vérification expérimentale.

Mais nous reviendrons sur ces questions au Chapitre suivant à propos de problèmes plus généraux que ceux que pose le jeu de pile ou face.

---

le calcul des intégrales J_n se déduit de la formule

$$\frac{d}{d\lambda}\left[\lambda^{n-1}e^{-\lambda^2}\right] = (n-1)\lambda^{n-2}e^{-\lambda^2} - 2\lambda^n e^{-\lambda^2}$$

qui intégrée entre les limites 0, $+\infty$ donne (on suppose $n - 1 > 0$) :

$$(n-1)J_{n-2} - 2J_n = 0.$$

D'autre part, nous avons trouvé directement

$$J_0 = \frac{1}{2}\sqrt{\pi}, \qquad J_1 = \frac{1}{2};$$

on calculera par la formule précédente, tous les J_n. On a, par exemple :

$$J_2 = \frac{1}{4}\sqrt{\pi}; \qquad J_3 = \frac{1}{8}\sqrt{\pi}; \qquad J_4 = \frac{3}{16}\sqrt{\pi}; \qquad J_5 = \frac{1\cdot 3\cdot 5\cdot 7}{32}\sqrt{\pi}.$$

CHAPITRE V

LOI DES GRANDS NOMBRES ET LOI DES ÉCARTS

28. Etude des épreuves répétées dans le cas général. — Nous allons étendre les résultats du chapitre précédent au cas où l'on considère, au lieu du jeu de pile ou face, une série d'épreuves dans laquelle la probabilité des deux alternatives a une valeur quelconque, supposée constante dans la série, mais non égale nécessairement à $\frac{1}{2}$. Nous désignerons par p la probabilité d'une des alternatives, que nous appelerons *alternative favorable*, pour fixer les idées ; la probabilité contraire sera $q = 1 - p$. Si l'on fait n épreuves consécutives, la probabilité pour que k d'entre elles soient favorables et $n - k$ défavorables est, si l'on specifie l'ordre dans lequel doivent se succéder les épreuves favorables et défavorables, égale à $p^k q^{n-k}$: si l'on ne specifie pas cet ordre, il faut multiplier ce nombre par le nombre C_n^k des combinaisons possibles : on obtient ainsi la valeur

$$P = \frac{n!}{k!\,(n-k)!}\,p^k q^{n-k}$$

Par exemple la probabilité d'amener le point 6 précisément deux fois en lançant un dé 3 fois de suite est

$$\frac{3!}{1!\,2!}\left(\frac{1}{6}\right)^2\frac{5}{6} = \frac{5}{72}$$

En effet, 3 combinaisons sont possibles, suivant que le coup de dés n'amenant pas 6 est le premier, le second ou le troisième et la probabilité de chacune de ces combinaisons est $\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)$.

Reprenons la valeur générale de P et cherchons à l'évaluer approximativement lorsque n et k sont très grands. Il est aisé de voir que P est maximum lorsque k et $n - k$ sont respectivement égaux à np et nq (en supposant ces nombres entiers : sinon il faudrait dire : égaux aux entiers les plus voisins de np et de nq). Nous poserons

$$k = np + t\sqrt{n}$$

et, par suite, puisque $p + q$ est égal à 1, nous aurons :

$$n - k = nq - t\sqrt{n}$$

nous aurons alors

$$\log P = \log(n!) - \log(np + t\sqrt{n})! - \log(nq - t\sqrt{n}) + k\log p + (n-k)\log q$$

et nous utiliserons la formule (2) du n° 20.

Nous remarquerons que l'on a

$$\log n! = \left(n + \frac{1}{2}\right)\log n - n + \log\sqrt{2\pi} + \frac{1+\theta_n}{12n}$$

et que l'on peut écrire

$$(n+1)\log n + k\log p + (n-k)\log q = \left(k + \frac{1}{2}\right)\log np + \left(n - k + \frac{1}{2}\right)\log nq - \frac{1}{2}\log pq$$

Il vient dès lors :

$$\log P = -\frac{1}{2}\log npq - \log\sqrt{2\pi} - \left(np + t\sqrt{n} + \frac{1}{2}\right)\log\left(1 + \frac{t}{p\sqrt{n}}\right) - \left(nq - t\sqrt{n} + \frac{1}{2}\right)\log\left(1 - \frac{t}{q\sqrt{n}}\right) + \frac{1+\theta_n}{12n} - \frac{1+\theta_k}{12k} - \frac{1+\theta_{n-k}}{12(n-k)}$$

En supposant $\frac{t}{p\sqrt{n}}$ inférieur à 1, on a

$$\left(np + t\sqrt{n} + \frac{1}{2}\right)\log\left(1 + \frac{t}{p\sqrt{n}}\right) - \left(np + t\sqrt{n} + \frac{1}{2}\right)\left(\frac{t}{p\sqrt{n}} - \frac{t^2}{2p^2n} + \frac{t^3}{3p^3n\sqrt{n}} \cdots\right)$$

$$\left(nq - t\sqrt{n} + \frac{1}{2}\right)\log\left(1 - \frac{t}{q\sqrt{n}}\right) - \left(nq - t\sqrt{n} + \frac{1}{2}\right)\left(-\frac{t}{q\sqrt{n}} - \frac{t^2}{2q^2n} - \frac{t^3}{3q^3n\sqrt{n}} \cdots\right)$$

On obtient ainsi aisément :

$$\log P = -\frac{1}{2}\log 2\pi npq - \frac{t^2}{2pq}$$

$$+ \frac{t^3}{6\sqrt{n}}\left(\frac{1}{p^2} - \frac{1}{q^2}\right) + \frac{t}{2\sqrt{n}}\left(\frac{1}{p} - \frac{1}{q}\right) + \cdots,$$

les termes non écrits étant beaucoup moins grands que les termes conservés.

Le plus important de ces termes conservés, qui disparaissait lorsque nous supposions $p = q$ est le suivant

$$\frac{t^3}{6\sqrt{n}}\left(\frac{1}{p^2} - \frac{1}{q^2}\right)$$

Mais il est aisé de voir, par une discussion dans le détail de laquelle il est inutile d'entrer, que ce terme est pratiquement négligeable, sauf dans le cas où $\frac{t}{p\sqrt{n}}$ ne serait pas petit [1].

En ne conservant que les termes principaux, on a

$$\log P = -\frac{1}{2}\log 2\pi npq - \frac{t^2}{2pq}$$

---

[1]. L'introduction de ces termes complémentaires a pour effet de remplacer la courbe symétrique des écarts par une *courbe dissymétrique* ; ces courbes dissymétriques ont été étudiées par divers mathématiciens et notamment par K. Pearson. Nous n'en parlerons pas ici.

nous poserons

$$t = \sqrt{2pq}\,\lambda$$

et λ sera l'écart relatif, l'écart absolu est $t\sqrt{n} = \lambda\sqrt{2npq}$ de sorte que *l'unité d'écart* est ici $\sqrt{2npq}$; avec les notations du chapitre précédent, nous avions :

$$n = 2m, \quad p = \frac{1}{2}, \quad q = \frac{1}{2}$$

et par suite

$$\sqrt{2npq} = \sqrt{m}$$

La probabilité P s'écrit

$$\mathrm{P} = \frac{1}{\sqrt{2\pi npq}}\, e^{-\lambda^2}$$

c'est la probabilité pour que l'écart relatif ait la valeur λ : la probabilité pour qu'il soit compris entre λ_1, et λ_2, supposés voisins, est

$$\frac{(\lambda_2 - \lambda_1)\sqrt{2npq}}{\sqrt{2\pi npq}}\, e^{-\lambda^2} = \frac{\lambda_2 - \lambda_1}{\sqrt{\pi}}\, e^{-\lambda^2}$$

La formule est exactement la même que dans le cas particulier examiné en premier lieu : la définition de l'unité d'écart a seule eu besoin d'être complétée : les conséquences sont donc les mêmes et il est inutile de les répéter à nouveau : nous allons insister sur quelques conséquences nouvelles, qui ont reçu le nom de *loi des grands nombres,* ou *théorème de Bernoulli.*

29. Loi des grands nombres. — Supposons que l'on effectue n épreuves successives, et désignons toujours par p la probabilité de l'alternative favorable et par q la probabilité de l'alternative contraire. Si nous désignons par λ l'écart relatif, le nombre des événements favorables sera, d'après la définition même de l'écart relatif, $np + \lambda\sqrt{2npq}$ et le nombre des événements contraires sera $nq - \lambda\sqrt{2npq}$; leur rapport sera

$$\frac{np + \lambda\sqrt{2npq}}{nq - \lambda\sqrt{2npq}}$$

calculons la différence entre ce rapport et le rapport $\frac{p}{q}$ des probabilités : on a

$$\frac{np + \lambda\sqrt{2npq}}{nq - \lambda\sqrt{2npq}} - \frac{p}{q} = \frac{\lambda\sqrt{2npq}\,(p+q)}{q\,(nq - \lambda\sqrt{2npq})}$$

En observant que l'on a $p + q = 1$ et en supprimant le facteur $\sqrt{n}$, cette différence prend la forme

$$\frac{\lambda\sqrt{2pq}}{q^2\sqrt{n} - \lambda q\sqrt{2pq}}$$

Le numérateur renferme le facteur λ et le dénominateur renferme, dans son terme principal le facteur $\sqrt{n}$: la différence calculée croît avec λ et décroît lorsque n croît. *Pour une valeur donnée de λ, elle tend vers zéro lorsque n augmente indéfiniment ;* donc, *si l'on suppose l'écart relatif λ inférieur à un nombre fixe le rapport entre le nombre des évènements favorables et le nombre des évènements contraires, tend vers le rapport $\frac{p}{q}$ de leurs probabilités, lorsque le nombre des épreuves augmente indéfiniment.* C'est la loi des grands nombres : on voit qu'elle se réduit, sous cette forme, à la simple constatation d'un fait analyttique.

30. Théorème de Bernoulli. — Si nous observons maintenant que la probabilité pour que la valeur absolue de λ dépasse un nombre A est $1 - \theta(A)$ et tend très rapidement vers zéro lorsque A augmente indéfiniment, on pourra donner ce nouvel énoncé.

Théorème de Jacques Bernoulli. — *Etant donné un nombre ε aussi petit que l'on veut, la probabilité pour que la différence entre le rapport observé du nombre des évènements favorables et le nombre des évènements contraires, d'une part, et le rapport théorique $\frac{p}{q}$, d'autre part, soit supérieure en valeur ab-*

solue à ε, *tend vers zéro lorsque le nombre* n *des épreuves augmente indéfiniment.*

Ecrivons, en effet, que la différence calculée plus haut est supérieure à ε, nous obtenons

$$\left| \frac{\lambda \sqrt{2pq}}{q^2 \sqrt{n} - \lambda q \sqrt{2pq}} \right| > \varepsilon$$

c'est-à-dire

$$\lambda > \frac{\varepsilon q^2 \sqrt{n}}{\sqrt{2pq} \pm \varepsilon q \sqrt{2pq}}$$

La probabilité pour que la relation précédente soit vérifiée est

$$1 - \theta\left(\frac{\varepsilon q^2 \sqrt{n}}{\sqrt{2pq} \pm \varepsilon q \sqrt{2pq}} \right)$$

Or, p et q sont fixes : donc, quelque petit que soit ε, le produit $\varepsilon \sqrt{n}$ augmente indéfiniment avec n et par suite la probabilité tend très rapidement vers zéro. Donnons quelques exemples.

31. Quelques problèmes. — Problème XVIII. *On jette un grand nombre de fois un même dé, la probabilité d'amener le point 6 est* $\frac{1}{6}$; *on calcule le rapport entre le nombre de fois où sort 6 et le nombre de fois où il ne sort pas ; au bout de combien d'épreuves la probabilité pour que ce rapport diffère de* $\frac{1}{5}$ *de moins de* $\frac{1}{1000}$ *est-elle inférieure à* $\frac{1}{10^8}$?

On a ici $p = \frac{1}{6}$ $q = \frac{5}{6}$; d'autre part, pour que $1 - \theta(\lambda)$ soit inférieur à $\frac{1}{10^8}$, λ doit dépasser 4 ; on doit donc avoir

$$\frac{4 \sqrt{\frac{1}{6} \cdot \frac{5}{6}}}{\left(\frac{5}{6}\right)^2 \sqrt{n} - 4 \frac{5}{6} \sqrt{2} - \frac{5}{6^2}} < \frac{1}{1000}$$

ce qui donne en négligeant le second terme du dénominateur,

$$n > \frac{10^2 . 5^2 . 5 . 6^2}{6^2 . 5^4} = \frac{10^2 . 5^2 . 6^2}{5^4}$$

Il suffit donc que l'on ait $n > 5\,000\,000$.

PROBLÈME XIX. *Paul joue contre Pierre à jeu non équitable ; les enjeux étant égaux, on suppose que la probabilité de gain pour Paul à chaque partie est* $\frac{1}{2} + \alpha$, *celle de Pierre étant par suite* $\frac{1}{2} - \alpha$. *Combien de parties doivent être jouées pour que la probabilité que Paul se retire du jeu avec une perte soit inférieure à* $\frac{1}{1\,000}$.

Nous supposerons que α est un nombre assez petit pour que nous puissions négliger son carré ; on a

$$\frac{p}{q} = \frac{\frac{1}{2} + \alpha}{\frac{1}{2} - \alpha} = \frac{1 + 2\alpha}{1 - 2\alpha} = 1 + 4\alpha + 8\alpha^2 \ldots$$

Nous prendrons $1 + 4\alpha$ pour valeur de $\frac{p}{q}$: pour que Paul se retire du jeu avec une perte il faut que le rapport entre le nombre des parties gagnées par lui et le nombre des parties perdues soit inférieur à 1, c'est-à-dire diffère du rapport théorique de plus de 4α, *dans un sens déterminé* ; pour que la probabilité de cet événement soit inférieure à $\frac{1}{1\,000}$, il faut que l'on ait

$$1 - \theta(\lambda) > \frac{1}{500}$$

$$\theta(\lambda) > 0{,}998$$

ce qui donne, en se bornant au chiffre des dixièmes

$$\lambda > 2{,}3$$

Nous ne ferons pas d'erreur sensible en remplaçant p et q par $\frac{1}{2}$ dans la formule n° 29 et en y négligeant le second terme du

dénominateur ; il vient alors

$$\frac{\lambda \frac{1}{\sqrt{2}}}{\frac{1}{4}\sqrt{n}} = 4\lambda$$

on en conclut, λ devant être supérieur à 2,2

$$n > \frac{(2.2)^2}{16\alpha^2}$$

Si l'on prend, par exemple $\alpha = \frac{1}{100}$, il vient

$$n > \frac{(220)^2}{16} = 55^2 = 3025$$

Si donc les probabilités de gain pour les joueurs sont 0,51 et 0,49, celui qui est avantagé par les conditions a 999 chances sur 1 000 de se retirer avec un gain, si l'on joue environ 3 000 parties.

Problème XX. *Appliquer les résultats précédents au jeu de rouge et noir à la roulette.* Sur 37 numéros, 18 sont rouges et 18 sont noirs ; le 37ème, qui est le zéro, n'est ni rouge ni noir. Lorsque le rouge sort, par exemple, le banquier reçoit les mises noires et double les mises rouges ; lorsque le zéro sort, le banquier retient la moitié de toutes les mises, noires ou rouges.

Considérons un joueur qui ponterait constamment sur *rouge* ; on peut dire que sa probabilité de gain est 18 sur 36,5 et sa probabilité de perte 18,5 sur 36,5 puisque le zéro n'entraîne qu'une demi perte (1). On a donc ici :

$$\frac{1}{2} + \alpha = \frac{18,5}{36,5}$$

$$\alpha = \frac{0,25}{36,5}$$

(1) On pourrait aussi remarquer que si le zéro sort, tout se passe comme si le banquier versait au ponte $\frac{1}{4}$ de l'enjeu et le ponte au banquier les $\frac{3}{4}$ de l'enjeu ; la probabilité du gain est donc 18,25 sur 37 et la probabilité de perte 18,75 sur 37. Cette manière de raisonner conduirait à des résultats sensiblement équivalents à ceux du texte ; en réalité ni l'une ni l'autre n'est absolument rigoureuse, car la convention particulière relative au zéro ne permet d'appliquer que d'une manière approchée les résultats du problème précédent.

et, par suite, la relation $n > \frac{(2,2)^2}{16\,\varepsilon^2}$ devient

$$n > (2,2)^2(36,5)^2 > (80,3)^2.$$

Il suffit donc que n dépasse 6 500 pour que le banquier ait plus de 999 chances sur 1 000 de ne pas perdre. En réalité, les chances de pertes du banquier sont encore plus faibles, car il n'a pas affaire à un joueur unique, mais à plusieurs dont les mises se répartissent sur rouge et sur noire : et il est clair que, si elles se répartissent également, sa certitude de ne pas perdre est absolue.

Problème XXI. *Quel est, au jeu non équitable défini au problème XIX le nombre de parties pour lequel la perte que Paul a une chance sur quatre de dépasser, est la plus élevée?* En supposant l'écart relatif λ de sens déterminé et égal à l'écart médian, nous nous trouverons bien dans les conditions de l'énoncé, car les 4 hypothèses suivantes

$$\begin{aligned} -\infty &< \lambda < -0,1769 \\ -0,1769 &< \lambda < 0 \\ 0 &< \lambda < 0,1769 \\ -0,1769 &< \lambda < \infty \end{aligned}$$

sont également probables : la probabilité de chacune d'elles est $\frac{1}{4}$. Pour cette valeur de λ, le nombre des parties gagnées par Paul est

$$\left(\frac{1}{2} - \varepsilon\right)n - \lambda\sqrt{2npq}$$

et le nombre de parties gagnées par Pierre est

$$\left(\frac{1}{2} - \varepsilon\right)n + \lambda\sqrt{2npq}.$$

La perte de Paul sera donc

$$2\lambda\sqrt{2npq} - 2\varepsilon n$$

à condition que cette quantité soit positive : lorsque n croît suffisamment, elle devient négative, c'est-à-dire que le jeu, en se

prolongeant, augmente les chances de gain pour Paul. Le maximum de la quantité précédente s'obtiendra en égalant a zéro sa dérivée par rapport à n : en prenant [1] $p = q = \frac{1}{2}$, cette quantité se réduit à

$$\lambda \sqrt{2n} - 2\alpha n.$$

On doit donc avoir

$$\frac{\lambda}{\sqrt{2\sqrt{n}}} - 2\alpha = 0$$

c'est-à-dire

$$n = \frac{\lambda^2}{8\alpha^2} = \frac{(0,48)^2}{8\alpha^2}$$

et la perte P de Paul est alors

$$P = \lambda \sqrt{2n} - 2\alpha n = \frac{\lambda^2}{4\alpha} = \frac{(0,48)^2}{4\alpha}.$$

En supposant, comme dans le problème XX, $\alpha = \frac{0,25}{36,5}$ et en remplaçant 0,48 par 0,5 et 36,5 par 36

$$n = \frac{36^2}{2} = 648$$

$$P = \frac{(0,5)^2}{4\alpha} = \frac{36}{4} = 9.$$

Ainsi, le nombre de parties le plus défavorable est ici 648 et la perte possible une fois sur 4 est égale a 9. Remarquons que l'unité d'écart est ici $\sqrt{2npq} = 18$: l'écart médian est environ 9. D'autre part, la probabilité de gain pour le banquier étant $\frac{1}{2} + \frac{1}{4 \times 36}$, le nombre des parties qu'il devrait théoriquement gagner est $324 + 4,5$; les quatre hypothèses suivantes sont donc également probables :

1° Le banquier gagne plus de $324 + 4,5 - 9 = 324 + 13,5$ parties ; son gain dépasse 27.

---

(1) Nous laissons au lecteur le soin de constater que cette simplification n'entraîne pas ici d'erreur sensible.

2° Le banquier gagne plus de $32\frac{1}{2} + 1,5$ parties, mais moins de $32\frac{1}{2} + 1,5 + 9$; son gain est compris entre 9 et 27.

3° Le banquier gagne moins de $32\frac{1}{2} + 1,5$ parties mais plus de $32\frac{1}{2} + 1,5 - 9$; il peut gagner ou perdre, mais dans les deux cas une somme inférieure à 9.

4° Le banquier gagne moins de $32\frac{1}{2} + 1,5 - 9$ parties ; sa perte est supérieure à 9.

On voit que les résultats d'ensemble, comme on devait s'y attendre, sont loin d'être défavorables au banquier, le jeu est toujours inéquitable en sa faveur ; le cas précédent est cependant celui où il est le plus exposé à une perte relativement considérable : il a une chance sur quatre de perdre au moins neuf fois la mise ; si cette mise est une fraction importante de sa fortune totale, il peut légitimement hésiter à accepter le jeu, bien qu'il soit dans l'ensemble avantageux pour lui. Au contraire si on lui propose de jouer 10 000 parties consécutives *et de ne régler les comptes qu'à la fin du jeu* (ou qu'il trouve à emprunter pour payer en cas de perte), il peut accepter ces conditions avec la certitude presque absolue d'un gain final (d'après les résultats du problème XX).

32. Cas où la probabilité a deux valeurs distinctes. — Supposons que les n épreuves considérées se décomposent en deux groupes, renfermant respectivement n_1 et n_2 épreuves et que la probabilité de l'événement favorable soit p_1 pour le premier groupe et p_2 pour le second groupe. Par exemple, on a deux urnes : l'une d'elles renferme a_1 boules blanches et b_1 boules noires et la seconde a_2 boules blanches et b_2 boules noires : on extrait n_1 boules de la première urne et n_2 de la seconde (en remettant chaque fois la boule extraite pour ne pas changer la composition des urnes) : on a ici $p_1 = \frac{a_1}{a_1 + b_1}$; $p_2 = \frac{a_2}{a_2 + b_2}$, si l'on regarde comme favorable l'extraction d'une boule blanche.

Si les extractions étaient exactement conformes aux probabilités, le nombre des événements favorables serait $n_1 p_1 + n_2 p_2$; en réalité il sera $n_1 p_1 + h_1$ pour le premier groupe et $n_2 p_2 + h_2$ pour le second, c'est à dire au total

$$n_1 p_1 + n_2 p_2 + h_1 + h_2 = n_1 p_1 + n_2 p_2 + h,$$

en désignant par h l'écart total : nous poserons

$$h_1 = \lambda_1 \sqrt{2n_1 p_1 q_1}$$
$$h_2 = \lambda_2 \sqrt{2n_2 p_2 q_2}$$

et nous rechercherons la probabilité pour que l'écart h ait une valeur déterminée, ou plutôt pour qu'il soit compris entre deux limites très rapprochées que nous appellerons h et $h + dh$. Ceci peut être réalisé avec une valeur quelconque x de l'écart h_1 ; il suffit que l'on ait

$$h < h_1 + h_2 = x + h_2 < h + dh$$

c'est-à-dire

$$h - x < h_2 < h - x + dh.$$

La probabilité pour que l'écart h_2 soit compris entre ces limites est, en posant

$$h - x = \lambda \sqrt{2n_2 p_2 q_2}$$

donnée par la formule

$$\frac{1}{\sqrt{\pi}} e^{-\lambda^2} d\lambda = \frac{1}{\sqrt{2\pi n_2 p_2 q_2}} e^{-\frac{(h-x)^2}{2n_2 p_2 q_2}} dh$$

Telle est la valeur de la probabilité correspondant à la valeur x de l'écart h_1 ; pour obtenir la probabilité totale il faut faire la somme des produits obtenus en multipliant ces probabilités partielles par la probabilité des diverses hypothèses possibles sur la valeur x ; or, la probabilité pour que x soit compris entre $x + dx$ est

$$\frac{1}{\sqrt{2\pi n_1 p_1 q_1}} e^{-\frac{x^2}{2n_1 p_1 q_1}} dx$$

La probabilité pour que h soit compris entre h et $h + dh$ est donc

$$\frac{dh}{\sqrt{2\pi n_1 p_1 q_1}\sqrt{2\pi n_2 p_2 q_2}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2n_1 p_1 q_1} - \frac{(h-x)^2}{2n_2 p_2 q_2}} dx.$$

Nous sommes amenés à calculer l'intégrale

$$\int_{-\infty}^{+\infty} e^{-(ax^2+bx+c)}\,dx.$$

Nous utiliserons l'identité

$$ax^2+bx+c = a\left(x+\frac{b}{2a}\right)^2+\frac{4ac-b^2}{4a}$$

et nous poserons

$$\sqrt{a}\left(x+\frac{b}{2a}\right)=y$$

$$\frac{4ac-b^2}{4a}=\rho$$

Il vient alors

$$\int_{-\infty}^{+\infty} e^{(ax^2+bx+c)}dx=\int_{-\infty}^{+\infty}\frac{e^{-y^2-\rho}dy}{\sqrt{a}}=\frac{e^{-\rho}}{\sqrt{a}}\int_{-\infty}^{+\infty}e^{-y^2}dy$$

$$=e^{-\rho}\frac{\sqrt{\pi}}{\sqrt{a}}.$$

Pour appliquer cette formule, nous poserons

$$2n_1p_1q_1=u^2_1$$

$$2n_2p_2q_2=u^2_2$$

en remarquant que u_1 et u_2 sont les unités d'écart correspondant aux deux groupes d'épreuves considérées isolément ; nous avons

$$\frac{x^2}{u^2_1}+\frac{(h-x)^2}{u^2_2}=ax^2+bx+c$$

c'est-à-dire

$$a=\frac{1}{u^2_1}+\frac{1}{u^2_2}\qquad b=\frac{-2h}{u^2_2}\qquad c=\frac{h^2}{u^2_2}$$

$$\rho=\frac{4ac-b^2}{4a}=\frac{h^2}{u^2_1+u^2_2}.$$

La valeur de la probabilité est donc

$$\frac{dh}{\pi u_1 u_2} e^{-\frac{h^2}{a}} \sqrt{\pi} \qquad \frac{dh}{\pi u_1 u_2} e^{-\frac{h^2}{u^2_1 + u^2_2}} \sqrt{\pi} \cdot \sqrt{\frac{1}{u^2_1} + \frac{1}{u^2_2}}$$

Or si l'on pose

$$u^2 = u^2_1 + u^2_2$$

elle prend la forme

$$\frac{1}{u\sqrt{\pi}} e^{-\frac{h^2}{u^2}} dh$$

et enfin, en écrivant

$$h = \lambda u$$

on retrouve la forme normale

$$\frac{1}{\sqrt{\pi}} e^{-\lambda^2} d\lambda.$$

La loi des écarts relatifs λ est donc la même que dans le cas d'un seul groupe d'épreuves, mais l'unité d'écart u est définie par la relation

$$u = \sqrt{u^2_1 + u^2_2} = \sqrt{2n_1 p_1 q_1 + 2n_2 p_2 q_2}.$$

33. Substitution d'un seul groupe d'épreuves à deux groupes. — Supposons que l'on détermine les nombres p et q par les relations

$$np = n_1 p_1 + n_2 p_2$$
$$nq = n_1 q_1 + n_2 q_2.$$

dans lesquelles n, n_1, n_2, p_1, q_1, p_2, q_2 ont les mêmes significations qu'au numéro précédent. Comme on a

$$p_1 + q_1 = 1, \qquad p_2 + q_2 = 1, \qquad n = n_1 + n_2$$

ces relations donnent

$$p + q = 1.$$

Si l'on fait n épreuves dans lesquelles la probabilité de l'événement favorable est p, les valeurs théoriques (ou valeurs les plus probables) de l'événement favorable et de l'événement contraire seront respectivement np et nq : elles sont donc les mêmes que dans les deux groupes d'épreuves précédemment envisagés. *Mais l'unité d'écart ne sera pas la même.* Nous désignerons par U cette unité d'écart correspondant aux n épreuves ; on a

$$U^2 = 2npq.$$

Nous avons trouvé, d'autre part, pour l'unité d'écart u dans le cas de deux groupes d'épreuves

$$u^2 = u_1^2 + u_2^2 = 2(n_1p_1q_1 + n_2p_2q_2)$$

On a donc :

$$U^2 - u^2 = 2(npq - n_1p_1q_1 - n_2p_2q_2)$$

Multiplions les deux membres de cette relation par $\frac{n}{2} = \frac{n_1 + n_2}{2}$; il vient

$$\frac{n(U^2 - u^2)}{2} = n^2pq - (n_1 + n_2)(n_1p_1q_1 + n_2p_2q_2).$$

Mais

$$n^2pq = np \cdot nq = (n_1p_1 + n_2q_2)(n_1q_1 + n_2q_2).$$

Donc il vient :

$$\frac{n(U^2 - u^2)}{2} = (n_1p_1 + n_2p_2)(n_1q_1 + n_2q_2) - (n_1 + n_2)(n_1p_1q_1 + n_2p_2q_2)$$
$$+ n_1n_2(p_1 - p_2)(q_2 - q_1).$$

Mais les relations

$$p_1 + q_1 = 1$$
$$p_2 + q_2 = 1$$

donnent

$$q_2 - q_1 = p_1 - p_2.$$

On a donc finalement

$$U^2 - u^2 = \frac{2n_1n_2}{n}(p_1 - p_2)^2.$$

On voit que si p_1 est différent de p_2, c'est-à-dire si les deux groupes sont réellement distincts, U^2 est toujours supérieur à u^2, c'est-à-dire que l'*unité d'écart est plus grande dans l'épreuve unique*. Cette règle se confirme immédiatement dans le cas particulier où l'on a :

$$p_1 = 1 \qquad q_1 = 0$$
$$p_2 = 0 \qquad q_2 = 1$$

c'est-à-dire ou, en reprenant l'exemple des urnes, on remplace n tirages dans une urne renfermant des boules blanches et noires en nombres proportionnels à p et q par $n_1 = np$ tirages dans une urne ne renfermant que des boules blanches et $n_2 = nq$ tirages dans une urne ne renfermant que des boules noires ; il est clair que dans cette seconde série d'épreuves, il n'y aura pas d'écarts.

Évaluons la variation relative de l'unité d'écart ; on a

$$\frac{U^2 - u^2}{U^2} = \frac{2n_1 n_2}{2n^2 pq}(p_1 - p_2)^2$$

Si l'on pose $n_1 = k_1 n$ d'où $n_2 = (1 - k_1)n$, il vient

$$\frac{U^2 - u^2}{U^2} = \frac{k_1(1 - k_1)}{p(1 - p)}(p_1 - p_2)^2.$$

Dans le cas où les nombres k_1 et p sont compris entre $\frac{1}{4}$ et $\frac{3}{4}$ on a

$$\frac{3}{16} < k_1(1 - k_1) < \frac{1}{4}$$
$$\frac{3}{16} < p(1 - p) < \frac{1}{4}$$

et la fraction $\frac{k_1(1 - k_1)}{p(1 - p)}$ est comprise entre $\frac{3}{4}$ et $\frac{4}{3}$; c'est donc le terme $(p_1 - p_2)^2$ qui fait connaître l'ordre de grandeur du rapport

$$\frac{U^2 - u^2}{U^2} = 1 - \frac{u^2}{U^2}.$$

Supposons, par exemple, que l'on ait

$$n_1 = n_2$$

$$p_1 = q_2 \qquad p_2 = q_1$$

et par suite $p = q = \frac{1}{2}$; il vient :

$$\frac{U^2 - u^2}{U^2} = (p_1 - p_2)^2 = (1 - 2p_1)^2$$

puisque $p_2 = q_1 = 1 - p_1$,

$$U^2 = \frac{u^2}{1 - (1 - 2p_1)^2} = \frac{u^2}{4p_1q_1}.$$

Si l'on suppose

$$p_1 = \frac{3}{4} \qquad p_2 = q_1 = \frac{1}{4}$$

il vient

$$U^2 = \frac{4u^2}{3}$$

$$U = \frac{2u}{\sqrt{3}}.$$

Ainsi, si au lieu de faire des tirages dans une urne unique renfermant autant de boules blanches que de boules noires, on fait le même nombre de tirages alternativement dans deux urnes dont la première renferme 3 fois plus de blanches que de noires, et la seconde 3 fois plus de noires que de blanches, l'unité d'écart dans le premier cas est égale à l'unité d'écart dans le second cas multiple par $\frac{2}{\sqrt{3}}$, c'est-à-dire par 1,15 environ.

Nous verrons quelle importance ont ces considérations lorsque nous étudierons les probabilités des causes.

34. Extension au cas de plusieurs groupes. — Nous n'entrerons pas dans le détail des développements analytiques par lesquels on peut traiter directement le cas de plusieurs groupes ; nous nous contenterons d'observer que l'on peut, pour réunir

plusieurs groupes en un seul, en réunir d'abord deux en un groupe unique, puis réunir le troisième à ce groupe unique, et ainsi de suite de proche en proche, de sorte que la règle générale se déduit de celle que nous avons établie dans le cas de deux groupes. Nous nous contenterons donc de l'énoncer

RÈGLE. — *Supposons que l'on effectue μ groupes d'épreuves, les nombres d'épreuves de ces groupes étant respectivement n_1, n_2, ... n_μ, les probabilités des événements favorables $p_1, p_2, \ldots, p_\mu$ et les probabilités contraires $q_1, q_2, \ldots, q_\mu$. Si l'on envisage cet ensemble d'épreuves, l'unité d'écart u sera donnée par la formule*

$$u^2 = u^2_1 + u^2_2 + \ldots + u_\mu{}^2$$

en posant

$$u^2_1 = 2n_1p_1q_1 \quad , \quad u^2_2 = 2n_2p_2q_2, \ldots, u_\mu{}^2 = 2n_\mu p_\mu q_\mu,$$

c'est-à-dire en désignant par u_1, u_2, ... u_μ les unités d'écarts qui correspondraient aux groupes considérés isolément.

Si l'on désigne par $n_1p_1 + n_2p_2 + \ldots + n_\mu p_\mu + \lambda u$ le nombre d'événements favorables, l'écart relatif λ a une probabilité normale, c'est-à-dire que la probabilité pour qu'il soit compris entre λ et $\lambda + d\lambda$ est

$$\frac{1}{\sqrt{\pi}} e^{-\lambda^2} d\lambda.$$

Cherchons l'écart U dans le cas où l'on remplacerait les μ groupes d'épreuves par une série unique de n épreuves équivalentes dans lesquelles les probabilités de l'événement favorable et de l'événement contraire seraient p et q. On aura donc

$$n = n_1 + n_2 + \ldots + n_\mu$$
$$np = n_1p_1 + n_2p_2 + \ldots + n_\mu p_\mu$$
$$nq = n_1q_1 + n_2q_2 + \ldots + n_\mu q_\mu,$$

et, par suite

$$n U^2 = 2n^2 pq = 2 \cdot np \cdot nq$$

$$\frac{n(U^2 - u^2)}{2} = \sum n_i p_i \sum n_i q_i - \sum n_i \sum n_i p_i q_i$$

$$= \sum\sum n_i n_k p_i q_k - \sum\sum n_i n_k p_i q_i$$

$$= \sum\sum n_i n_k (p_i - p_k)(q_i - q_k)$$

le signe $\sum$ indiquant que la dernière somme ne s'étend qu'une seule fois à chaque produit $n_i n_k$ (on ne considère pas $n_k n_i$, c'est-à-dire que l'on suppose, par exemple i supérieur à k),

On a donc finalement

$$U^2 = u^2 + \frac{2}{n} \sum\sum n_i n_k (p_i - p_k)^2.$$

Le second membre est essentiellement positif, comme on pouvait s'y attendre.

35. Résumé des résultats obtenus. — En résumé, l'étude d'un problème de probabilité discontinue, lorsque le nombre des épreuves est trop grand pour permettre l'énumération complète, exige :

1° La connaissance de la loi de probabilité des écarts relatifs λ. Cette loi est donnée par la fonction $\theta(\lambda)$.

2° Le calcul de l'unité d'écart u,

La probabilité pour qu'un écart soit compris entre $\lambda_1 u$ et $\lambda_2 u$ est dès lors égale à la probabilité pour que λ soit compris entre λ_1 et λ_2, c'est-à-dire, en supposant λ_1 et λ_2 positifs et $\lambda_1 < \lambda_2$, à :

$$\frac{1}{2}\left[\theta(\lambda_2) - \theta(\lambda_1)\right]$$

ce qui peut s'écrire

$$\frac{1}{2\sqrt{\pi}} \cdot \int_{\lambda_1}^{\lambda_2} e^{-\lambda^2} d\lambda$$

Or, en posant

$$\lambda u = h \quad , \quad \lambda_1 u = h_1 \quad , \quad \lambda_2 u = h$$

cette intégrale devient :

$$\frac{1}{2u\sqrt{\pi}} \int_{h_1}^{h_2} e^{-\frac{h^2}{u^2}} dh$$

Telle est la probabilité pour que l'écart absolu h soit compris entre h_1 et h_2 quels que soient leurs signes, pourvu que $h_2 - h_1$ soit positif. Dans le cas d'une série unique d'épreuves, on a

$$u^2 = 2npq$$

et la formule devient

$$\frac{1}{2\sqrt{2\pi npq}} \int_{h_1}^{h_2} e^{-\frac{h^2}{2npq}} dh.$$

Mais comme, pour le calcul numérique de ces expressions, on est forcé de recourir à la table des valeurs de $\theta\,\lambda$, *il est préférable de ne pas introduire l'écart absolu h et de s'en tenir aux formules faisant connaître la probabilité de l'écart relatif* λ *et la valeur de l'unité d'écart u.*

LIVRE II

PROBABILITÉS CONTINUES

CHAPITRE VI

DÉFINITION DE LA PROBABILITÉ GÉOMÉTRIQUE

36. Position d'un point sur un segment de droite. — Considérons un segment de droite AB et un point M que nous supposons *assujetti à rester sur* AB *fig.* 1. Quelle est la probabilité pour que le point M occupe sur AB une position déterminée? Cette question se distingue de celles que nous avons étudiées au livre I en ce que *le nombre des cas possibles est indéfini*, car le point M peut varier d'une manière continue entre A et B : il est donc nécessaire de donner une définition nouvelle de la probabilité [1]. Une telle définition est évidemment une *convention* pour le mathématicien, car celui-ci peut raisonner sur n'importe quelle définition : mais ce n'est pas une *convention arbitraire* ; elle est suggérée *à priori* par l'étude pratique de diverses questions et vérifiée *a posteriori* par la conformité de ses conséquences avec les observations. Quant le mathématicien dit que c'est une convention, cela veut dire simplement que, n'étant pas en état de *démontrer cette définition* d'une manière rigoureuse, il préfère la considérer comme une convention et rendre ainsi

---

(1) On doit faire observer que nous étudions l'hypothèse où l'ensemble des cas possibles est *fini* et l'hypothèse où cet ensemble a, dans la terminologie de M. Cantor la *puissance du continu*, mais non pas l'hypothèse où cet ensemble est infini, mais *dénombrable* : par exemple, où la variable peut prendre toutes les valeurs entières ; ou toutes les valeurs rationnelles, ou algébriques.

Cette dernière hypothèse n'a guère été étudiée jusqu'ici ; elle mériterait de l'être, surtout au point de vue mathématique. Voir Borel : Les probabilités dénombrables et leurs applications arithmétiques (Rendiconti del Circolo matematico di Palermo, 1909).

ses raisonnements abstraits absolument inattaquables. La difficulté qui consiste à légitimer la convention se trouve ainsi transportée au début de l'étude de chaque cas particulier concret.

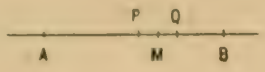

Fig. 1.

La définition que nous adopterons est la suivante.

Définition. — *La probabilité pour que le point* M *se trouve sur un certain segment* PQ *de* AB *est proportionnelle à la longueur de ce segment* (1).

Conséquence.— Si on suppose M assujetti à se trouver sur AB, la probabilité pour que M soit sur AB est 1 : donc *la probabilité pour que* M *soit sur* PQ *est égale au rapport* $\frac{PQ}{AB}$.

Il importe de bien se rendre compte que cette définition et ses conséquences ne sont pas sans pouvoir présenter certaines difficultés : on pourrait les exposer sous forme géométrique, en faisant usage par exemple de la transformation homographique, ou projection conique : nous préférons donner une exposition algébrique.

37. Objection à la définition. — Un nombre x est assujetti à être compris entre 5 et 10 ; la probabilité pour qu'il soit compris entre 5 et 6 sera, d'après ce qui précède :

$$\frac{6-5}{10-5}=\frac{1}{5}.$$

D'autre part, considérons le carré de x ; si x est compris entre 5 et 10, x^2 est compris entre 25 et 100 ; si x est compris entre 5 et 6, x^2 est compris entre 25 et 36 ; or la probabilité pour que

(1) Nous considérons ici et dans ce qui suit les segments au point de vue de leur mesure arithmétique, et non de leur valeur algébrique, c'est-à-dire que nous ne tenons pas compte de leurs signes. M. Jules Tannery m'a fait observer, il y a déjà longtemps, qu'il serait plus logique d'appeler *vecteur* un segment dirigé, le mot segment n'impliquant aucune idée de direction.

x^2 soit compris entre 25 et 36, sachant que x^2 est assujetti à être compris entre 25 et 100, est

$$\frac{36-25}{100-25}=\frac{11}{75}.$$

Considérons d'autre part $\frac{1}{x}$; si x est compris entre 5 et 10, ou entre 5 et 6, $\frac{1}{x}$ sera compris entre $\frac{1}{5}$ et $\frac{1}{10}$ ou entre $\frac{1}{5}$ et $\frac{1}{6}$. La probabilité pour que $\frac{1}{x}$ soit compris entre $\frac{1}{5}$ et $\frac{1}{6}$, sachant qu'il doit être compris entre $\frac{1}{5}$ et $\frac{1}{10}$, est

$$\frac{\frac{1}{5}-\frac{1}{6}}{\frac{1}{5}-\frac{1}{10}}=\frac{\frac{1}{30}}{\frac{1}{10}}=\frac{1}{3}$$

On conclut de là que la définition est inacceptable, car, dit-on, il est absurde que les résultats soient différents lorsque l'on remplace x par une fonction simple de x; d'une manière plus générale, $f(x)$ désignant une fonction quelconque de x que l'on supposera cependant varier toujours dans le même sens : constamment croissante, ou constamment décroissante, et x étant assujetti à se trouver compris entre a et b, la probabilité pour que x soit compris entre c et d (on suppose $a<c<d<b$), sera

$$P=\frac{f(d)-f(c)}{f(b)-f(a)}$$

car c'est la probabilité pour que $f(x)$ soit compris entre $f(c)$ et $f(d)$. En choisissant convenablement $f(x)$, on peut donner à P une valeur *absolument quelconque*.

Telle est l'objection, que nous avons tenu à exposer avec toute sa force, mais dont l'importance pratique n'est pas très grande.

Nous reviendrons plus loin sur l'intérêt qu'elle présente au point de vue théorique, à propos des profondes remarques de M. Poincaré sur le fait que le résultat final est, dans certaines questions, indépendant du choix des fonctions arbitraires, nous restons pour l'instant sur le terrain concret.

A ce point de vue pratique, l'objection précédente doit simplement et ce n'est pas là une utilité négligeable *attirer notre attention sur les erreurs que pourrait entraîner un mauvais choix de la variable indépendante*. Mais, en fait, ce choix est presque toujours imposé d'une manière évidente par les conditions même de la question posée, lorsqu'il s'agit d'une question concrète et non d'une question abstraite, et c'est une simple plaisanterie que de prétendre modifier arbitrairement ce choix par un changement de variable qui n'est qu'un artifice analytique, sans rapport avec la réalité. Nous nous rendrons un compte exact de la manière dont la variable indépendante doit être divisée en étudiant des problèmes particuliers (voir le Chapitre VII ; mais ce choix n'a un sens que si l'on se borne aux problèmes dont l'énoncé est assez précis pour qu'une vérification expérimentale du résultat puisse être tout au moins tentée (voir les remarques faites au début du chapitre VII [1]).

38. Position d'un point dans le plan ou dans l'espace. — Si un point M doit se trouver dans une aire plane A, la probabilité pour que M se trouve dans une portion déterminée S de cette aire est *proportionnelle à la surface de* S, elle est donc égale

Fig. 2.

au rapport $\frac{S}{A}$, en désignant par S et par A les superficies des aires S et A (*fig.* 2). Analytiquement cette probabilité s'exprime

---

[1] Le lecteur qui s'intéresse surtout aux applications peut ajourner la lecture de la fin de ce chapitre et passer à l'étude des problèmes. (Ch. VII).

par le quotient de deux intégrales doubles

$$\frac{\iint_{A} dxdy}{\iint_{S} dxdy}.$$

Si le point M doit se trouver dans l'espace à l'intérieur d'un volume V, la probabilité pour qu'il se trouve dans une portion U de ce volume sera égale au rapport $\frac{U}{V}$ et se calculera par le rapport des deux intégrales

$$\frac{\iiint_{U} dxdydz}{\iiint_{V} dxdydz}$$

Ces expressions permettent de calculer la valeur probable ou valeur moyenne d'une fonction quelconque $f(x, y, z)$: c'est la somme des produits que l'on obtient en multipliant chaque valeur possible de cette fonction par sa probabilité, c'est-à-dire

$$\frac{\iiint_{V} f(x, y, z)\, dxdydz}{\iiint_{V} dxdydy}$$

En employant le langage de la géométrie à n dimensions, on peut étendre ces définitions et ces résultats au cas où le nombre des variables indépendantes dépasse 3 ; cette considération est souvent utile en physique mathématique (voir le chapitre VIII).

39. Autres probabilités géométriques. — On considère parfois des probabilités qui ne sont pas relatives à la position d'un point, mais d'un autre élément géométrique, dépendant d'une ou de plusieurs variables ; il est des cas où le choix des variables est très simple et s'impose d'après les données du problème ; il en est d'autres où il est plus compliqué ; c'est en traitant des problèmes particuliers (chapitre VII) que l'on peut le

mieux se rendre compte des difficultés qui se présentent et de la marche à suivre pour les surmonter : donnons cependant quelques indications générales.

Citons d'abord le cas où l'on considère un point assujetti à rester sur une courbe dans le plan ou sur une surface dans l'espace : les cas les plus fréquents sont ceux où il s'agit d'un cercle ou d'une sphère : il est alors naturel de considérer la probabilité comme proportionnelle à la longueur de l'arc de cercle ou à l'aire de la portion de surface sphérique.

Un cas absolument analogue au précédent est celui d'une droite assujettie à passer par un point fixe dans le plan ou dans l'espace : on considérera son intersection avec un cercle ou avec une sphère ayant pour centre le point fixe. Dans le plan, la position de la droite sera fixée par l'angle φ qu'elle fait avec une direction fixe ; cet angle φ varie de 0 à π s'il s'agit d'une droite indéfinie dans les deux sens et de 0 à 2π s'il s'agit d'une *demi-droite*, limitée au point fixe O. Bornons-nous à ce dernier cas. La probabilité pour que l'angle fait par la demi-droite avec la direction fixe soit compris entre φ_1 et φ_2 (on suppose $0 < \varphi_1 < \varphi_2 < 2\pi$) est alors

$$\frac{\varphi_2 - \varphi_1}{2\pi} = \frac{\int_{\varphi_1}^{\varphi_2} d\varphi}{\int_0^{2\pi} d\varphi}$$

et la valeur probable ou moyenne d'une fonction $f(\varphi)$ est

$$\frac{1}{2\pi} \int_0^{2\pi} f(\varphi)\, d\varphi$$

Pour fixer la position d'un point sur la sphère (ou la direction d'une demi-droite dans l'espace) on fera usage de deux coordonnées : on peut choisir la *longitude* θ variant de $-\pi$ à $+\pi$ et la latitude φ variant de $-\frac{\pi}{2}$ à $+\frac{\pi}{2}$; les coordonnées d'un point de la sphère sont alors

$$\begin{cases} X = R \cos\theta \cos\varphi \\ Y = R \sin\theta \cos\varphi \\ Z = R \sin\varphi \end{cases}$$

L'angle θ varie de $-\pi$ à $+\pi$ et l'angle φ de $-\frac{\pi}{2}$ à $+\frac{\pi}{2}$. L'élément de surface de la sphère a pour expression

$$R^2 \cos\varphi d\theta d\varphi$$

de sorte que la surface totale est

$$\int_{-\pi}^{+\pi} d\theta \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} R^2 \cos\varphi d\varphi = 4\pi R^2$$

La probabilité pour que la direction de la demi droite soit comprise entre certaines limites est donc

$$\frac{1}{4\pi} \iint_{(s)} \cos\varphi d\theta d\varphi$$

l'intégrale double étant étendu au champ s de variation des angles θ et φ. La valeur moyenne d'une fonction $f(\theta, \varphi)$ est de même

$$\frac{1}{4\pi} \int_{-\pi}^{+\pi} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} f(\theta, \varphi) \cos\varphi d\theta d\varphi$$

Les cas précédents sont ceux qui se présentent le plus fréquemment ; citons seulement pour mémoire les probabilités relatives à la position dans le plan d'une droite indéfinie, dépendant d'un ou deux paramètres, ou d'un segment limité de ligne droite ou de ligne courbe. Nous nous bornerons à traiter quelques exemples particuliers (Problème XXIX, n° **46**).

CHAPITRE VII

QUELQUES PROBLÈMES DE PROBABILITÉS GÉOMÉTRIQUES

40. Remarque préliminaire. — Nous laissons de côté maintenant les considérations générales développées dans le chapitre précédent pour nous consacrer à l'étude de problèmes *concrets* de probabilités géométriques. Par problèmes concrets, nous entendons ceux dont l'énoncé est suffisamment précis pour que l'on en puisse déduire une méthode *de vérification expérimentale* des résultats. Il paraîtra sans doute singulier à quelques-uns qu'il puisse être question d'une *vérification expérimentale* dans laquelle intervient le hasard : si l'on veut vérifier expérimentalement les résultats obtenus sur la probabilité au jeu de pile ou face, ne pourra-t-il pas arriver que l'on tombe sur une série extraordinaire de 1 000 parties consécutives donnant pile ? Devra-t-on en conclure que la théorie est fausse ? Non certainement, mais on recommencera l'expérience, et un nombre suffisant d'expériences donnera un résultat cadrant avec la théorie. Un événement tel que l'arrivée 1 000 fois de suite de pile au jeu de pile ou face, est aussi invraisemblable que le suivant : un bon physicien, avec d'excellents instruments, tente par une méthode nouvelle la mesure de l'accélération due à la pesanteur et trouve un nombre très différent du nombre actuellement admis ; la première idée qui lui viendra sera de recommencer ses expériences, et c'est seulement après de très nombreux essais qu'il

ajoutera quelque foi à son résultat et cherchera la cause du désaccord ailleurs que dans une erreur fortuite [1].

De même, ce n'est pas par une vérification unique, c'est-à-dire par une seule série d'expériences, que l'on doit chercher à contrôler les valeurs d'une probabilité géométrique : on fera de nombreuses séries et si, sauf quelque série exceptionnelle, elles donnent des résultats voisins entre eux et voisins du résultat théorique, ce sera une forte présomption en faveur de celui-ci. On pourra même acquérir une certitude sur le point suivant : si l'emploi de deux méthodes différentes conduisent, dans l'étude théorique du problème, à deux résultats très différents entre eux, l'expérience permet d'éliminer avec sécurité l'un de ces résultats et par suite de condamner l'une des deux méthodes.

41. Une difficulté de la vérification expérimentale. — Dans la vérification expérimentale des résultats relatifs aux probabilités géométriques, il se présente une difficulté sur laquelle il n'est pas inutile d'insister, car elle est d'un ordre assez général et tient, dans certains cas, à la nature profonde des choses.

Considérons le cas le plus simple où il s'agit de déterminer les probabilités pour qu'un point M soit situé sur une certaine portion PQ d'un segment AB (*fig.* 3). Si l'on veut calculer expéri-

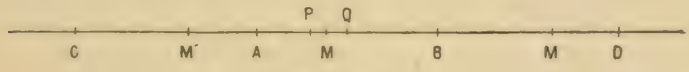

Fig. 3.

mentalement cette probabilité, on est amené à se donner un segment matérialisé AB et à chercher à désigner *au hasard* un point M de ce segment. Le segment AB pourra être, par exemple, une règle de bois, sur laquelle on jettera « au hasard » un objet mé-

[1] Un sceptique dira : vous ne tenez donc compte de l'expérience que si elle confirme les prévisions de la théorie ? Nullement, mais la conformité aux résultats d'expériences antérieures est une raison pour supposer qu'il ne s'est pas glissé quelqu'une de ces erreurs grossières qui sont le résultat de coïncidences infiniment rares quand l'expérimentateur est bon et les instruments bien vérifiés. De telles coïncidences sont aussi rares qu'une série très anormale au jeu de pile ou face.

tallique de très petites dimensions, qui figurera le point M. La difficulté qui se présente alors est la suivante : si l'on veut que, à chaque expérience, le point M soit *certainement* sur AB, on sera forcément conduit à « viser » plus ou moins grossièrement la partie médiane de AB et, sans qu'il soit besoin d'aucune théorie, le simple bon sens indique que la probabilité pour que M soit sur un segment PQ, de longueur déterminée, sera plus élevée lorsque PQ sera vers le milieu de AB et plus faible lorsque PQ sera vers les extrémités. On pourrait, par les méthodes de la théorie des erreurs, traiter le problème en se plaçant à ce point de vue : mais ce serait là un problème compliqué, dont l'utilité n'est pas en rapport avec la complication. *Nous laisserons complètement ce point de vue de côté*, c'est-à-dire que nous ne supposerons pas que l'on cherche à placer M sur AB : il arrivera alors fréquemment que M sera sur le prolongement de AB, à droite ou à gauche, en M′ ou en M″. On peut convenir de ne pas tenir compte de ces positions de M ; on peut aussi considérer les segments CA et BD égaux à AB et situés à droite et à gauche de AB (*fig.* 3) et convenir d'assimiler chaque position M de M sur CA, par exemple, à la position *homologue* sur AB, c'est-à-dire à la position que vient occuper le point M″ lorsque l'on amène par une translation, CA sur AB (C venant en A et A en B). Un peu de réflexion montre que cette manière de procéder n'est légitime que si les segments égaux à AB sont très nombreux dans le champ de l'expérience, c'est-à-dire si chacun d'eux est petit par rapport aux variations de la position de M : on aura, dans ce cas, une série de segments égaux (*fig.* 4) : A_4A_3, A_3A_2, A_2A_1, A_1A, AB, BB_1, etc., considérés tous comme homologues et

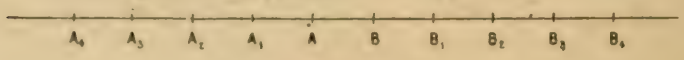

Fig. 4.

l'on considérera chaque position de M sur l'un d'eux comme équivalente à la position homologue sur AB. Cette forme spéciale donnée à l'expérience n'est jamais indispensable. Mais elle peut être commode dans certains cas (voir le problème de l'aiguille, p. 104), ce qui est *essentiel*, c'est de ne pas faire l'expé-

rience en s'efforçant de placer le plus souvent possible M sur AB, car on fausse alors les conditions.

42. Problèmes relatifs à la position de points sur une droite. — Problème XXII. *Deux points M et M′ sont assujettis à se trouver sur un segment AB, de longueur l ; quelle est la probabilité pour que la distance MM′ soit inférieure à kl, k étant donné inférieur à un.*

Désignons par x la distance AM ; la probabilité pour que cette distance soit comprise entre x et $x + dx$ est $\frac{dx}{l}$. Prenons (*fig.* 5) de part et d'autre du point M deux longueurs, MC = MD $= kl$; pour que MM′ soit inférieur à kl il faut et il suffit que M

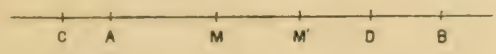

Fig. 5.

soit compris entre C et D ; mais M′ doit aussi être compris entre A et B ; il faut donc distinguer plusieurs cas suivant les positions relatives des points C, A, D, B ; comme le segment CD est égal à $2kl$, il sera inférieur ou supérieur à AB suivant que k sera inférieur ou supérieur à $\frac{1}{2}$; nous examinerons successivement ces deux cas, à chacun desquels correspondent 3 cas différents de figure

$k < \frac{1}{2}$

Ordre des points	Inégalité pour x	Longueur du segment sur lequel peut se trouver M′
CADB (*fig.* 6)	$0 < x < kl$	$AD = x + kl$

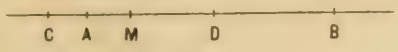

Fig. 6.

ACDB (*fig.* 7)	$kl < x < (1 - k)\,l$	$CD = 2kl$

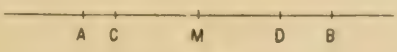

Fig. 7.

ACBD (*fig.* 8)	$(1 - k)\,l < x < l$	$CB = kl + l - x$

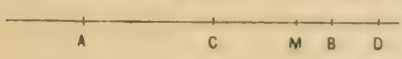

Fig. 8.

Cas où $k < \frac{1}{2}$	Inégalité pour x	Longueur du segment sur lequel peut se trouver M
CADB	$o < x < (1-k)$	$AD = x + kl$
CABD	$(1-k)\,l < x < kl$	$AB = l$
ACBD	$kl < x < l$	$CB = kl + l - x$

La probabilité cherchée est donc, pour $k < \frac{1}{2}$:

$$\frac{1}{l}\left[\int_0^{kl}(x+kl)\frac{dx}{l}+\int_{kl}^{(1-kl)}2kl\frac{dx}{l}+\int_{(1-kl)}^{l}(kl+l-x)\frac{dx}{l}\right]$$

et, pour $k > \frac{1}{2}$,

$$\frac{1}{l}\left[\int_0^{(1-kl)}(x+kl)\frac{dx}{l}+\int_{(1-kl)}^{kl}l\,\frac{dx}{l}\int_{kl}^{l}(kl+l-x)\frac{dx}{l}\right].$$

Le calcul est facilité si l'on remarque que, pour des raisons de symétrie évidente dans chacune des formules, la troisième intégrale est égale à la première.

On trouve ainsi, pour $k < \frac{1}{2}$

$$2\left[\frac{k^2}{2}+k^2\right]+2k(1-2k)=2k-k^2$$

et pour $k > \frac{1}{2}$

$$2\left[\frac{(1-k)^2}{2}+k(1-k)\right]+2k-1=2k-k^2.$$

Le résultat est le même dans les deux hypothèses ; on s'en rend compte aisément en remplaçant le calcul des intégrales par les évaluations de surfaces ; si l'on représente le segment $AB = l$ (*fig.* 9 et 10) et que l'on élève en chaque point M de AB une ordonnée MQ égale à la longueur du segment que peut occuper M lorsque M est au pied de l'ordonnée, on obtient une aire couverte de hachures, dont le quotient par l^2 représente la

probabilité cherchée. Dans la figure 9, correspondant à $k < \frac{1}{2}$, on a

$$AA' = lk, \quad AH = HL = kl \quad PL = 2kl \quad LN = l(1 - 2k).$$

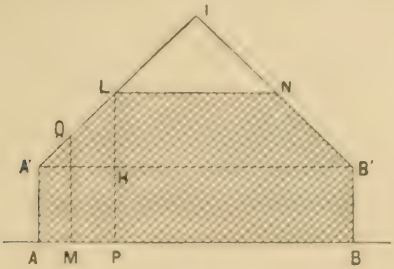

Fig. 9.

Dans la figure 10, correspondant à $k > \frac{1}{2}$, on a

$$AA' = kl \quad A'H = HL = (1 - k)l \quad IL = l \quad LN = l(2k - 1)$$

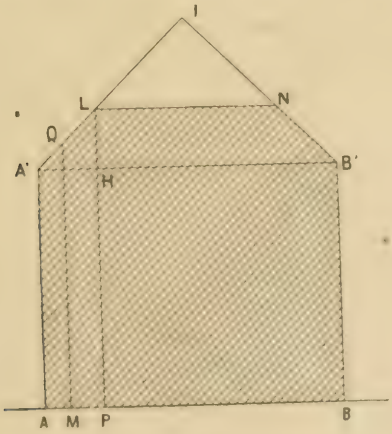

Fig. 10.

Sur les deux figures, l'aire couverte de hachures est égale à rectangle AA'B'B + triangle A'B'I — triangle LNI, c'est-à-dire à :

$$AB \times AA' + \frac{1}{4}\overline{A'B'}^2 - \frac{1}{4}\overline{LN}^2.$$

Mais, dans les deux figures, on a

$$\overline{AB} \times \overline{AA'} = l^2 k$$
$$\overline{AB}^2 = l^2$$
$$\overline{LN}^2 = l^2 (1 - 2k)^2$$

car LN est égal à $\pm l(1 - 2k)$; la probabilité qui s'obtient en divisant l'aire par l^2, est donc

$$k + \frac{1 - (1 - 2k)^2}{4} = 2k - k^2.$$

Indiquons encore une autre méthode pour arriver au même résultat ; posons

$$AM = x \qquad AM' = y.$$

La probabilité pour que AM soit compris entre $x + dx$ et M' entre $y + dy$ est égale à $\frac{dx\,dy}{l^2}$. Si donc on représente le point P dont les coordonnées rectangulaires sont x et y (*fig.* 11), le point sera intérieur à un carré OABC de côté l, et la probabilité pour

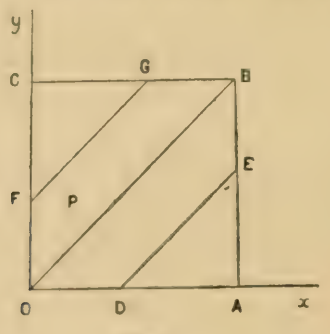

Fig. 11.

qu'il soit dans une aire donnée est proportionnelle à cette aire. Or la condition imposée par l'énoncé est que l'on ait

$$|x - y| < kl$$

ce qui exige que P soit compris entre les droites DE et FG, parallèles à OB et ayant pour équations respectives

$$x - y = \pm kl.$$

La probabilité cherchée est donc égale au rapport de l'aire ODEBGFO à l'aire du carré ; or les deux triangles rectangles ADE et CFG forment, par leur réunion, un carré de côté $(1-k)l$ car on a $OD = OF = kl$; la probabilité cherchée est donc

$$\frac{l^2 - (1-k)^2 l^2}{l^2} = 2k - k^2.$$

Comme il fallait s'y attendre la probabilité croît constamment lorsque k varie de 0 à 1 ; pour $k = 0$ elle est égale à 0 et pour $k = 1$ elle est égale à 1 ; pour $k = \frac{1}{2}$, elle est égale à $\frac{3}{4}$.

Problème XXIII. *On prend au hasard deux points* M *et* M′ *sur un segment* AB, *de longueur* l ; *quelle est la valeur probable de la distance* MM′ ?. *Quelle est la valeur probable de son carré* ?

On peut déduire la solution de ce problème du résultat que nous venons de trouver ; nous savons que la probabilité pour que MM′ soit inférieur à kl est égale à $2k - k^2$; la probabilité pour que MM′ soit compris entre kl et $(k + dk)\,l$ est égale à la différentielle de $2k - k^2$ c'est-à-dire à $2(1-k)\,dk$; la valeur probable de MM′ est par suite

$$\int_0^1 2(1-k)\,kl\,dk = \frac{l}{3}.$$

La valeur probable de MM′ est donc $\frac{l}{3}$ ou $\frac{AB}{3}$; la valeur probable de son carré est

$$\int_0^1 2(1-k)\,k^2 l^2 dk = \frac{l^2}{6}$$

c'est-à-dire $\frac{\overline{AB}^2}{6}$ elle est supérieure au carré de la valeur probable de MM′.

Traitons la question directement ; si l'on pose $AM = x$, $BM' = y$, la valeur probable de MM′ est

$$\frac{1}{l^2}\int_0^l \int_0^l |x - y|\,dxdy$$

car $MM' = |x - y|$. Or, on a évidemment

$$\int_0^l |x-y|\,dx = \int_0^y (y-x)\,dx + \int_y^l (x-y)\,dx$$
$$= \left[xy - \frac{x^2}{2}\right]_0^y + \left[\frac{x^2}{2} - xy\right]_y^l$$
$$= \frac{l^2}{2} - ly + y^2$$

et ensuite

$$\int_0^l \left(\frac{l^2}{2} - ly + y^2\right) dy = \left[\frac{l^2 y}{2} - \frac{ly^2}{2} + \frac{y^3}{3}\right]_0^l = \frac{l^3}{3}.$$

La valeur probable est donc $\frac{l}{3}$.

Quant à la valeur probable de $\overline{MM'}^2$, elle est

$$\frac{1}{l^2}\int_0^l\int_0^l (x-y)^2\,dx\,dy = \frac{1}{l^2}\int_0^l \frac{(l-y)^3 + y^3}{3}\,dy$$
$$= \frac{1}{l^2}\left[\frac{y^4 - (l-y)^4}{12}\right]_0^l = \frac{l^2}{6}.$$

De même, la valeur probable de $\overline{MM'}^{2n}$ serait

$$\frac{1}{l^2}\int_0^l\int_0^l (x-y)^{2n}\,dx\,dy = \frac{1}{l^2}\int_0^l \frac{(l-y)^{2n+1} + y^{2n+1}}{2n+1}\,dy$$
$$= \frac{1}{l^2}\left[\frac{y^{2n+2} - (l-y)^{2n+2}}{(2n+1)(2n+2)}\right]_0^l = \frac{2l^n}{(2n+1)(2n+2)}$$

43. Problèmes relatifs à la position de points sur un cercle. — Problème XXIV. *On marque au hasard deux points* M *et* M′ *sur un cercle ; quelle est la probabilité pour que le plus petit arc* MM *soit inférieur à* α. On suppose, bien entendu $\alpha < \pi$. Ce problème est plus simple que le précédent car on peut prendre M quelconque : la probabilité n'est pas changée : M′ doit se trouver sur l'arc 2α dont le milieu est en M ; la probabilité cherchée est donc

$$\frac{2\alpha}{2\pi} = \frac{\alpha}{\pi}.$$

La probabilité pour que MM' soit compris entre α et $\alpha + d\alpha$ est donc $\frac{d\alpha}{\pi}$ et la valeur probable de MM' est

$$\int_0^{\pi} \frac{\alpha d\alpha}{\pi} = \frac{\pi}{2}.$$

Problème XXV. *On marque au hasard n points sur un cercle* $M_1, M_2, \ldots M_n$. *Quelle est l'espérance mathématique d'un joueur à qui on promettrait autant de francs qu'il y a d'arcs* M_iM_k *inférieurs à* α ?

Nous venons de voir que la probabilité pour que l'un des axes soit inférieur à α est $\frac{\alpha}{\pi}$; or s'il y a n points, le nombre des combinaisons 2 à 2 telles que M_iM_k est $\frac{n(n-1)}{2}$; l'espérance mathématique est donc

$$\frac{n(n-1)}{2}\frac{\alpha}{\pi}.$$

Si l'on suppose que α ait pour mesure 1° et que l'on ait $n = 25$, cette expression devient

$$\frac{25 \cdot 24}{2}\frac{1}{180} = \frac{5}{3}.$$

L'espérance mathématique est sensiblement supérieure à un : la *probabilité* pour qu'il y ait au moins un arc inférieur à 1° est au contraire sensiblement inférieure à 1 ; mais l'espérance mathématique s'accroît par suite de la possibilité pour le joueur de toucher une somme relativement considérable (qui peut atteindre jusqu'à 300 francs), dans le cas où de nombreux arcs M_iM_k sont inférieurs à 1°.

44. Problème relatif à la position de points dans une aire plane. — Problème XXVI. *On prend au hasard deux points* M *et* M' *à l'intérieur d'un carré de côté a ; quelle est la valeur probable du carré de la distance* MM' ?

Si nous désignons par x, y les coordonnées de M et par x', y' les coordonnées de M', nous sommes conduits à calculer l'inté-

grale quadruple

$$\frac{1}{a^4}\int_0^a\int_0^a\int_0^a\int_0^a\left[(x-x')^2+(y-y')^2\right]dxdydx'dy'.$$

Or, il vient successivement

$$\int_0^a\left[(x-x')^2+(y-y')^2\right]dx=\frac{(a-x')^3+x'^3}{3}+a(y-y')^2$$

$$\int_0^a\left[\frac{(a-x')^3+x'^3}{3}+a(y-y')^2\right]dx'=\frac{a^4}{6}+a^2(y-y')^2$$

$$\int_0^a\int_0^a\left[\frac{a^4}{6}+a^2(y-y')^2\right]dydy'=\frac{a^6}{6}+\frac{a^6}{6}=\frac{a^6}{3}$$

La valeur moyenne cherchée est donc $\frac{a^2}{3}$.

45. Problèmes relatifs à la position de points sur la sphère. — Ces problèmes ont une certaine importance, car on peut y rattacher, comme nous le verrons au Livre III, diverses questions relatives à la position des étoiles sur la sphère céleste.

Problème XXVII. *Deux points* M *et* M′ *sont pris au hasard sur la surface de la sphère ; quelle est la probabilité pour que le plus petit arc de grand cercle* MM′ *soit inférieur à* α.

La probabilité sera la même quelle que soit la position de M ; or lorsque le point M est fixé, M′ doit se trouver sur une calotte sphérique AB, correspondant à un demi angle au centre $MOA=\alpha$ (*fig.* 12). On a donc, en désignant par R le rayon de la sphère

$$MP=OM-OP=R(1-\cos\alpha).$$

Le rapport de l'aire de la calotte à l'aire de la sphère est

$$\frac{MP}{2R}=\frac{1-\cos\alpha}{2}=\sin^2\frac{\alpha}{2}.$$

Telle est la probabilité cherchée ; si α est très petit, on peut remplacer le sinus par l'arc et prendre pour valeur approchée $\frac{\alpha^2}{4}$.

Bertrand indique en même temps que la méthode précédente, un autre raisonnement qui conduit à un résultat tout différent [1]. Lorsqu'on donne deux points M et M' l'arc de grand cercle qui les joint est déterminé [2] ; tous les arcs de grand cercle étant

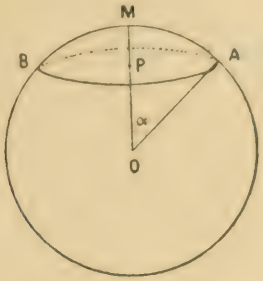

Fig. 12.

analogues sur la sphère, on ne change pas la probabilité en fixant cet arc de grand cercle : or la probabilité pour que deux points M et M' d'un cercle soient tels que l'arc MM' soit inférieur à α est, comme nous l'avons vu $\frac{\alpha}{\pi}$; ce résultat est très différent du précédent, surtout si α est très petit : si la mesure de α est 1°, on a $\alpha = \frac{\pi}{180}$ et par suite ;

$$\frac{\alpha^2}{4} = \frac{\pi^2}{360^2} \qquad \frac{\alpha}{\pi} = \frac{1}{180}.$$

Le rapport de la seconde valeur à la première est $\frac{720}{\pi}$ c'est-à-dire plus de 200. Doit-on en conclure avec Bertrand, que le problème proposé ne peut pas être résolu et que la première solution que nous en avons donnée est incorrecte ? Cette solution est, au contraire, la seule correcte, si l'on admet le postulat relatif à la probabilité élémentaire, c'est-à-dire si l'on considère toutes les portions égales de la sphère comme équivalentes entre elles, au point de vue de la probabilité qu'a le point M ou le point M' de s'y trouver. Or, dans son second raisonnement, Bertrand paraît

---

[1] Calcul des probabilités. Chapitre I.

[2] On peut écarter le cas dont la probabilité est évidemment nulle, où M et M' sont diamétralement opposés.

bien admettre cette sorte d'*homogénéité*, si l'on peut ainsi dire, qui fait regarder tous les points de la sphère comme équivalents entre eux. Ce second raisonnement renferme donc une cause d'inexactitude, qu'il n'est pas inutile de mettre en évidence. Examinons le d'un peu près.

On commence par affirmer que la probabilité n'est pas changée lorsque l'on fixe l'arc de grand cercle sur lequel se trouve M et M′ et la position de M sur ce grand cercle : cela est évident par raison de symétrie. L'erreur commence lorsque, le point M et grand cercle étant fixes, on considère que la probabilité pour que M′ se trouve sur un arc donné de ce grand cercle est proportionnelle à la longueur de cet arc. Si l'arc de grand cercle est sans épaisseur il faut, pour parler rigoureusement, assigner la valeur zéro à la probabilité pour que M et M′ soient sur ce cercle : on doit, pour ne pas avoir ce facteur zéro qui rend tout calcul impossible, considérer un faisceau délié d'arcs de grand cercle passant par un même point M et il est alors visible que la probabilité est plus grande pour M′ situé à un quadrant de M que situé au voisinage de M (*fig.* 13). Supposons, par exemple

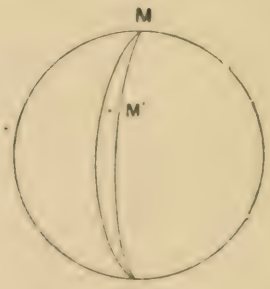

Fig. 13.

que la sphère soit la sphère terrestre, M le pôle nord et M′ le point de chute d'un bolide qui se meut dans l'espace suivant une loi inconnue, avec une vitesse beaucoup plus grande que la vitesse de la terre. On demande quelle est la probabilité pour que le point de chute de M′ ait une latitude nord supérieure à 89° ; le raisonnement de Bertrand revient à dire : on peut supposer connue la longitude du point de chute M′, supposer, par exemple que M′ se trouve sur le méridien de Paris ; tous les points de ce

méridien sont alors également probables. Nous allons voir que, même en acceptant ce point de vue, on arrivera, en serrant de près la vérification expérimentale, à constater que tous les points du méridien ne sont pas également probables : comment, en effet, constatera-t-on qu'un point M est sur le méridien de Paris? En déterminant sa longitude à l'aide d'observations astronomiques et d'un chronomètre ; cette détermination se fait avec une certaine précision angulaire : supposons que cette précision soit de $0^s,1$, par exemple : cela veut dire que l'on considère, non pas la ligne idéale sans épaisseur que serait le méridien théorique, mais l'espace compris entre $0^s,1$ de longitude est de $0^s,1$ de longitude ouest, et cet espace est beaucoup plus grand à l'équateur, qu'au voisinage du pôle : la probabilité pour que la latitude de M soit comprise entre 0° et 1° est donc beaucoup plus grande que la probabilité pour que cette latitude soit comprise entre 89° et 90°.

On pourrait penser que, pour réfuter le raisonnement précédent, il suffirait de substituer aux observations astronomiques pour la détermination de la longitude, des mesures géodésiques. Supposons donc que, par des mesures d'une grande précision, on soit arrivé à tracer un méridien sur la surface terrestre ; on aura du moins fixé un certain nombre de repères, suffisamment rapprochés pour que deux au moins soient aperçus de tout point du méridien et, sur chacun de ces repères un trait vertical extrêmement délié fait connaître son intersection avec le plan du méridien. On supposera alors que le centre de gravité M du bolide se trouve sur le méridien ainsi défini : la probabilité ne sera-t-elle pas la même à l'équateur qu'au pôle? Oui, sans doute, si les traits verticaux tracés sur les repères ont partout la même épaisseur, mais cette épaisseur uniforme des traits est contradictoire avec la notion du méridien : quelle que soit l'épaisseur choisie pour le trait du repère situé à l'équateur (disons, par exemple, un dixième de millimètre), les deux bords de ce trait définissent avec le pôle *deux méridiens*, dont l'angle est sans doute très petit, puisque leur écartement à l'équateur est seulement un dixième de millimètre, mais cet écartement est encore plus petit lorsqu'on s'approche du pôle ; si l'on plaçait à l'équateur deux traits verticaux voisins, ayant chacun un dixième de millimètre d'épaisseur, mais le bord droit de l'un coïncidant avec le bord

gauche de l'autre, il ne serait pas possible de prolonger jusqu'au pôle les méridiens ainsi définis, *en leur conservant leur largeur de un dixième de millimètre* : si on parvenait à réaliser cette construction géodésique infiniment délicate, les traits d'un dixième de millimètre d'épaisseur qui seraient tracés à une latitude quelconque *empiéteraient l'un sur l'autre*, d'autant plus que la latitude s'approcherait davantage de 90°, de sorte que, dans la région commune à ces traits, on n'aurait aucune raison pour dire que l'on se trouve sur l'un ou sur l'autre des méridiens définis à l'équateur.

Cette discussion un peu longue n'était pas inutile pour réduire à sa juste valeur la boutade de Bertrand, à laquelle il n'attachait pas lui-même plus de valeur qu'il ne convient [1]. On doit la considérer comme un exemple mettant en garde contre des raisonnements inexacts.

Problème XXVIII. *On place n points au hasard sur une sphère ; quelle est l'espérance mathématique d'une personne qui recevrait* 1 *franc pour chaque couple de points dont la distance angulaire est inférieure à* α ?

La probabilité pour chaque couple est, si α est petit, sensiblement égale à $\frac{\alpha^2}{4}$; l'espérance mathématique est donc

$$\frac{n(n-1)}{2} \cdot \frac{\alpha^2}{4} = \frac{n(n-1)\alpha^2}{8}$$

résultat important dans l'étude de la probabilité des causes. (voir n° **70**).

46 Autres problèmes. — Problème XXIX. Problème de l'aiguille. — *Ayant tracé sur une feuille de papier horizontale des parallèles équidistantes, on jette au hasard une aiguille parfaitement cylindrique : quelle est la probabilité pour que l'aiguille rencontre l'une des parallèles ?*

---

(1) Voir Darboux, Eloge historique de Joseph Bertrand. A propos d'un paradoxe analogue, dont nous parlerons tout à l'heure, M. Darboux dit que Bertrand connaissait la solution, mais préférait la laisser deviner à son lecteur.

Ce problème est célèbre ; il se prête aisément à une vérification expérimentale, qui a été souvent tentée, car il est connu depuis longtemps (1).

Désignons par $2a$ la distance des parallèles AB et CD et par $2l$ la longueur de l'aiguille PQ. On peut supposer que le milieu M de l'aiguille est sur une certaine perpendiculaire EF aux parallèles AB et CD. On désignera par x la longueur EM, supposée inférieure à FM et l'on supposera $l < a$, c'est-à-dire que l'aiguille ne pourra rencontrer qu'une seule au plus des parallèles. Nous verrons, en effet, que, si elle peut en rencontrer plusieurs, il faut, pour arriver à un résultat simple, substituer à la recherche de la probabilité celle de l'espérance mathématique d'une personne à qui seraient promis autant de francs qu'il y a de points d'intersection.

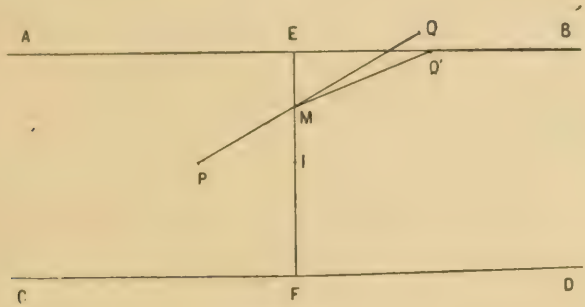

Fig. 14.

Pour que l'aiguille puisse couper AB, il est nécessaire que x soit inférieur à l : si l'on désigne par I le milieu de EF il est clair que la probabilité pour que M étant sur EI, EM soit compris entre x et $x + dx$, est $\frac{dx}{a}$. Le point M étant fixé, il faut que l'angle aigu EMQ soit inférieur à l'angle EMQ′ obtenu en supposant l'extrémité Q′ sur AB ; on a

$$\cos \widehat{EMQ'} = \frac{EM}{MQ'}$$

c'est-à-dire

$$\widehat{EMQ'} = \text{arc cos } \frac{x}{l}$$

(1) C'est Buffon qui en a donné le premier la solution correcte.

La probabilité pour que l'angle aigu EMQ $\left(\text{compris entre } 0 \text{ et } \frac{\pi}{2}\right)$ soit inférieur à $\widehat{EMQ'}$ est

$$\frac{2}{\pi} \arccos \frac{x}{l}$$

et la probabilité pour que l'aiguille rencontre AB est par suite

$$\frac{2}{\pi} \int_0^{a} \arccos \frac{x}{l} \frac{dx}{a}$$

Si l'on pose $x = ly$ cette intégrale devient

$$\frac{2l}{a\pi} \int_0^{1} \arccos y \, dy$$

La probabilité cherchée est proportionnelle à $\frac{2l}{a\pi}$; on calcule l'intégrale en posant $y = \cos t$ et l'on a

$$\int_0^{1} \arccos y \, dy = \int_{\frac{\pi}{2}}^{0} -t \sin t \, dt = (t \cos t - \sin t)_{\frac{\pi}{2}}^{0} = 1$$

Si l'on suppose, en particulier $2l = a$, c'est-à-dire la longueur de l'aiguille égale à la moitié de la distance entre les parallèles, la probabilité devient égale à $\frac{1}{\pi}$; si donc on divise le nombre des épreuves par le nombre des rencontres, on obtiendra approximativement la valeur de π. La vérification a été tentée et a donné des résultats parfaitement satisfaisants, aussi satisfaisants que le serait le calcul du rapport entre le nombre des extractions et le nombre des boules blanches extraites d'une urne qui renfermerait $31\,4\,159\,265$ boules dont $21\,4\,159\,265$ boules noires et $100\,000\,000$ boules blanches. Sur $4\,000$ tirages, l'unité d'écart s'obtiendrait en prenant

$$n = 4000 \qquad p = \frac{1}{\pi} \qquad q = \frac{\pi - 1}{\pi}$$

d'où

$$\sqrt{2npq} = \sqrt{\frac{8000\,(\pi - 1)}{\pi}}$$

c'est-à-dire un peu plus de 10 : en supposant un écart égal à 10 l'erreur commise sur $\frac{1}{\pi}$ serait environ $\frac{1}{300}$ et l'erreur commise sur π environ $\frac{1}{100}$ c'est-à-dire que la première décimale serait exacte et la seconde approchée à une unité près. Telle est l'approximation que l'on peut espérer d'une série de 5000 expériences sur l'aiguille : un tel résultat, répété plusieurs fois, suffit pour confirmer l'exactitude des principes invoqués et pour prouver l'inexactitude d'une méthode qui conduirait à un résultat sensiblement différent de $\frac{1}{\pi}$.

Un raisonnement ingénieux, dû à M. Barbier, [1] permet de calculer sans peine l'espérance mathématique du joueur qui doit recevoir 1 franc pour chaque point d'intersection de l'aiguille avec les parallèles : dans le cas $l < a$, comme il ne peut y avoir plus d'un de ces points, l'espérance mathématique se confond avec la probabilité.

Considérons une ligne polygonale (ouverte ou fermée) supposée matérialisée et jetons la au hasard sur les parallèles : il est clair que l'espérance mathématique du joueur qui doit recevoir 1 fr. pour chaque point d'intersection est égale à la somme des espérances mathématiques pour les divers côtés des polygones, lesquelles sont égales elles-mêmes à la somme des espérances analogues pour les segments égaux en lesquels on peut diviser les côtés (supposés commensurables par le mode de raisonnement classique). On en conclut que *l'espérance mathématique cherchée est proportionnelle à la longueur du polygone*. Or, si celui-ci se réduit à un cercle de diamètre $2a$, elle est évidemment égale à 2, car il y a toujours 2 points d'intersection et 2 seulement : la longueur d'un tel cercle est $2\pi a$: l'espérance mathématique *par unité de longueur* est donc

$$\frac{2}{2\pi a} = \frac{1}{\pi a}$$

Pour une aiguille de longueur $2l$ on retrouve le résultat $\frac{2l}{\pi a}$.

---

(1) *Journal de Liouville*, 1860 (2e série t. V.)

Pour bien montrer sur un exemple la différence entre l'espérance mathématique et la probabilité, cherchons, dans le cas où $2l$ dépasse $2a$, quelle est la probabilité pour que l'aiguille rencontre au moins une parallèle. Si nous revenons à la figure 14 nous remarquerons que, si l'aiguille rencontre au moins une parallèle, elle rencontre sûrement la plus voisine de son milieu; on peut donc supposer M entre E et I et on trouve en raisonnant comme plus haut, l'expression suivante de la probabilité.

$$\frac{2}{a\pi}\int_0^a \arccos\frac{x}{l}\,dx$$

Si l'on pose $x = l\cos t$ d'où $dx = -l\sin t\,dt$ et que l'on pose $a = l\cos\alpha$, il vient pour valeur de l'intégrale

$$\frac{2l}{a\pi}\int_\alpha^{\frac{\pi}{2}} t\sin t\,dt = \frac{2l}{a\pi}\left(\sin t - t\cos t\right)_\alpha^{\frac{\pi}{2}} = \frac{2l}{a\pi}\left(1 - \sin\alpha + \alpha\cos\alpha\right)$$

On pourrait vérifier que cette probabilité est inférieure à l'unité : nous nous bornerons à le constater en supposant l très grand par rapport à a et calculant dans ce cas sa valeur asymptotique ; nous avons posé ;

$$\cos\alpha = \frac{a}{l}$$

on en conclut

$$\sin\alpha = \left(1 - \frac{a^2}{l^2}\right)^{\frac{1}{2}} = 1 - \frac{a^2}{2l^2} + \dots$$

et, d'autre part $\frac{\pi}{2} - \alpha$ étant très petit est égal, aux termes du troisième ordre près à $\sin\left(\frac{\pi}{2} - \alpha\right)$ ou $\cos\alpha$; on a donc

$$\alpha = \frac{\pi}{2} - \frac{a}{l} + \dots$$

L'expression de la probabilité devient alors

$$\frac{2l}{a\pi}\left[\left(1 - \frac{a^2}{2l^2} + \dots\right) + \frac{a}{l}\left(\frac{\pi}{2} - \frac{a}{l} + \dots\right)\right]$$

c'est-à-dire

$$\frac{2l}{a\pi}\left[\frac{a\pi}{2l} - \frac{a^2}{2l^2} + \dots\right] = 1 - \frac{a}{l\pi} + \dots$$

Aussi lorsque l est grand par rapport à a, la probabilité *pour qu'il n'y ait aucune rencontre* est sensiblement égale à $\frac{a}{l\pi}$.

On peut établir ce résultat directement en observant que si l est grand par rapport à a, la probabilité pour que l'aiguille coupe une parallèle est sensiblement indépendante de la position P d'une de ses extrémités ; si l'on suppose P fixé, l'autre extrémité Q se trouve sur une circonférence de centre P et de rayon $2l$, circonférence dont la longueur est $4\pi l$; pour que l'aiguille ne coupe aucune parallèle, il faut que Q se trouve sur l'un des deux arcs RS, R'S' de cette circonférence qui sont compris entre les deux parallèles entre lesquelles se trouve P (*fig.* 15) ;

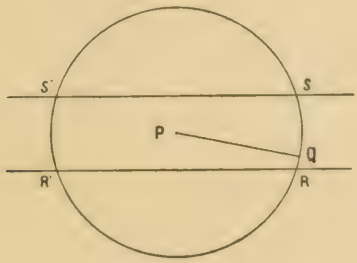

Fig. 15.

or, si le rayon du cercle est assez grand par rapport à la distance des parallèles, la longueur de chacun de ces arcs diffère très peu de la distance $2a$ de ces parallèles ; la somme de leurs longueurs est voisine de $4a$ et la probabilité pour que Q, qui doit se trouver sur la circonférence de longueur $4\pi l$, soit sur l'un de ces arcs, est

$$\frac{4a}{4\pi l} = \frac{a}{\pi l}$$

Notre raisonnement prouve que cette valeur est approchée par défaut, car la vraie longueur des arcs est sûrement *supérieure* à $2a$.

Problème XXX. *On trace une corde au hasard dans un cercle ; quelle est la probabilité pour que sa longueur soit supérieure au côté du triangle équilatéral inscrit ?*

Ce probleme est cité par Bertrand comme exemple d'un énoncé incomplet : il en donne en effet trois solutions qui conduisent à des résultats différents.

Première solution. — On peut, pour des raisons de symétrie, se donner la direction de la corde ; le point d'intersection de cette corde avec le diamètre perpendiculaire à cette direction devra alors se trouver sur un segment égal à la moitié de la longueur de ce diamètre (car la distance au centre du côté du triangle équilatéral inscrit est égale à la moitié du rayon) ; la probabilité est donc $\frac{1}{2}$.

Deuxième solution. — On peut, pour des raisons de symétrie, se donner une des extrémités de la corde sur le cercle ; la tangente en ce point et les deux côtés du triangle équilatéral inscrit ayant ce point pour sommet forment trois angles de 60° ; la direction de la corde doit être à l'intérieur de l'un de ces trois angles à l'exclusion des deux autres ; la probabilité est donc $\frac{1}{3}$.

Troisième solution. — Pour fixer la position d'une corde, il suffit de donner son milieu ; pour que la corde satisfasse à la condition de l'énoncé, il faut que son milieu soit intérieur à un cercle concentrique au cercle donné et de rayon moitié : la surface de ce cercle est le quart de la surface du cercle donné : la probabilité est donc $\frac{1}{4}$.

Doit-on penser que ces trois solutions sont également bonnes et, par suite, également mauvaises ? Nullement, il s'agit simplement de préciser le mode d'après lequel se fera la vérification expérimentale, c'est-à-dire comment on s'y prendra pour tracer une corde *au hasard* dans un cercle : si l'on assujettit cette corde à passer par un point fixe du cercle, ou si l'on fixe son milieu *au hasard*, c'est la seconde ou la troisième solution qui est la bonne : mais il est aisé de voir que la plupart des procédés naturels que l'on peut imaginer conduisent à la première.

Si par exemple, on jette un disque circulaire sur un plan sur

lequel sont tracées des droites, la probabilité pour que l'une des cordes interceptées soit supérieure au côté du triangle équilatéral est égale à $\frac{1}{2}$; il en est de même si l'on considère les cordes interceptées sur le disque circulaire de la lune par la trajectoire d'une étoile occultée, ou les cordes décrites dans le champ circulaire d'une lunette astronomique par des astres que l'on n'a pas visés, et qui occupent par suite une position quelconque dans le champ de la lunette.

Tout ceci apparaîtra plus clairement encore dans l'étude du problème suivant, généralisation de celui que nous venons de traiter.

Problème XXXI. *Quelle est la probabilité pour que la longueur d'une corde dans le cercle du rayon* R *soit comprise entre a et b ?*

Si l'on se donne la direction du diamètre perpendiculaire à la corde, on voit que la distance au centre est comprise entre $\sqrt{R^2 - \frac{a^2}{4}}$ et $\sqrt{R^2 - \frac{b^2}{4}}$: la probabilité cherchée est donc, en supposant $a < b$.

$$\frac{1}{R}\left[\sqrt{R^2 - \frac{a^2}{4}} - \sqrt{R^2 - \frac{b^2}{4}}\right]$$

La probabilité pour que la longueur soit comprise entre a et $a + da$ est par suite

$$(1) \qquad -\frac{1}{R}d\left(\sqrt{R^2 - \frac{a^2}{4}}\right) = \frac{a\,da}{2R\sqrt{4R^2 - a^2}}$$

La longueur *probable* d'une corde dans un cercle de rayon R est par suite :

$$\int_0^{2R} \frac{a^2 da}{2R\sqrt{4R^2 - a^2}}$$

Cette intégrale se calcule par la substitution

$$a = 2R\cos\varphi.$$

Elle devient

$$\int_0^{\frac{\pi}{2}} 2R\cos^2\varphi\, d\varphi$$

$$= \int_0^{\frac{\pi}{2}} R(1+\cos 2\varphi)\, d\varphi = \frac{R\pi}{2}.$$

La longueur probable d'une corde d'un cercle est donc égale à la longueur d'un quart de cercle.

On peut appliquer le résultat précédent à la solution du problème de l'aiguille. Nous conserverons les notations du Problème XXIX et nous bornerons au cas où $a = l$. Soient AB et CD deux parallèles voisines (*fig.* 16), EF la perpendiculaire com-

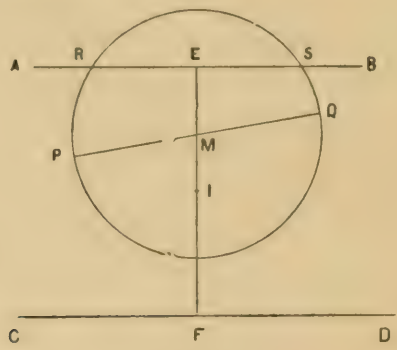

Fig. 16.

mune à AB et CD, I le milieu de EF ; on se bornera au cas où le milieu M de l'aiguille PQ est situé sur le segment EI et on cherchera la probabilité pour que, dans ces conditions, l'aiguille coupe AB. Décrivons le cercle de centre M et de rayon $MQ = l$; soient R et S les points d'intersection de ce cercle avec AB ; la probabilité cherchée est égale au rapport de l'arc RS à la demi-circonférence ; si l'on désigne par $2c$ la longueur de la corde RS, ce rapport est égal à

$$\frac{2}{\pi}\operatorname{arc\,sin}\frac{c}{l}$$

si donc on désigne par $\varphi(c)\,dc$ la probabilité pour que la lon-

gueur d'une corde, dans le cercle de rayon l, soit comprise entre $2c$ et $2c + 2dc$, la probabilité totale cherchée est

$$\int_o^l \frac{2}{\pi} \arcsin \frac{c}{l} \varphi(c)\, dc.$$

Or on a, d'après ce qui précède, (en posant dans la formule (1) : $a = 2c$, $\mathrm{R} = l$,

$$\varphi(c)dc = \frac{4cdc}{2l\sqrt{4l^2 - 4c^2}} = \frac{cdc}{l\sqrt{\mathrm{R}^2 - c^2}}$$

et, en posant $c = l \sin\theta$, l'intégrale devient

$$\int_o^{\frac{\pi}{2}} \frac{2}{\pi} \sin\theta d\theta = \frac{\pi}{2}$$

ce qui est bien le résultat déjà trouvé. Si l'on admettait la seconde ou la troisième des solutions proposées par Bertrand pour le problème XXX, il en résulterait une autre valeur pour $\varphi(c)$ et par suite une solution différente du problème de l'aiguille. Comme, pour celui-ci, la vérification expérimentale est facile, on met ainsi bien en évidence le caractère arbitraire de ces solutions ; elles correspondent à des hypothèses très spéciales et rarement réalisées sur la manière de faire l'expérience.

CHAPITRE VIII

INTRODUCTION DES FONCTIONS ARBITRAIRES

47. Équations générales. — Lorsque l'on ne fait aucune hypothèse sur la vérification expérimentale et que l'on se place à un point de vue purement abstrait, on peut introduire, comme l'a fait systématiquement M. Poincaré, une fonction arbitraire dans la définition de la probabilité.

Supposons tout d'abord qu'il s'agisse de la position d'un point sur une droite, position qui est fixée pour l'abscisse x. On désignera par $\varphi(x)\,dx$ la probabilité par que l'abscisse soit comprise entre x et $x + dx$. La seule condition à laquelle doive satisfaire a priori la fonction $\varphi(x)$ est d'être *positive;* mais le théorème des probabilités totales permet sans peine d'obtenir une seconde condition. Le point étant assujetti à être sur la droite, x a sûrement une valeur comprise entre $-\infty$ et $+\infty$; autrement dit la probabilité pour que x soit compris entre $-\infty$ et $+\infty$ est égale à 1 ; on a donc

$$\int_{-\infty}^{+\infty} \varphi(x)\,dx = 1. \tag{1}$$

Dans le cas où l'on sait que x est assujetti à être compris entre a et b on pourrait remplacer l'équation (1) par la suivante :

$$\int_{a}^{b} \varphi(x)\,dx = 1. \tag{2}$$

Mais on peut observer que l'équation (1) comprend l'équation (2) comme cas particulier ; il suffit de supposer que l'on a

$\varphi(x) = 0$ lorsque x est compris entre $-\infty$ et a, ou entre b et $+\infty$: l'équation (1) se réduit alors à l'équation (2).

Il est clair que par l'introduction de la fonction arbitraire φx, on supprime toutes les difficultés relatives aux changements de variables ; ou, du moins, car un artifice analytique ne saurait supprimer une difficulté réelle, on ajourne l'examen de ces difficultés au moment où l'on étudiera chaque application particulière et la théorie n'a pas à s'en embarrasser : elle s'établit sur la fonction arbitraire φx, qu'on est amené à remplacer dans chaque problème concret par une fonction bien déterminée, dépendant des conditions de ce problème.

La probabilité pour que x soit compris entre deux nombres c et d s'exprime par le quotient

$$\frac{\int_c^d \varphi(x)\,dx}{\int_{-\infty}^{+\infty} \varphi(x)\,dx} = \int_c^d \varphi(x)\,dx.$$

La première expression est souvent commode, car elle permet de remplacer, sans erreur, φx par une quantité proportionnelle. De même, si la position d'un point dépend de deux variables, on introduira une fonction arbitraire $\varphi x, y$ assujettie à la double condition d'être positive et de vérifier la relation

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \varphi(x, y)\,dx\,dy = 1$$

et la condition pour que le point $x\,y$ soit dans une aire (S) s'exprimera par l'intégrale double

$$\iint_{(S)} \varphi(x, y)\,dx\,dy.$$

Si l'on remplace x et y par les coordonnées polaires ρ et ω, au moyen des formules

$$x = \rho \cos \omega$$
$$y = \rho \sin \omega$$

on sait que l'on doit dans l'intégrale double remplacer $dx\,dy$ par $\rho\,d\rho\,d\omega$, car on a

$$\frac{D(x,y)}{D(\rho,\omega)} = \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \omega} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \omega} \end{vmatrix} = \rho.$$

On en conclut que la fonction $\varphi(x, y)$ doit être remplacée par

$$\Phi(\rho, \omega) = \rho\varphi(\rho\cos\omega, \rho\sin\omega).$$

D'une manière générale si l'on pose :

$$x = f(\alpha,\beta)$$
$$y = g(\alpha,\beta)$$

et si l'on pose

$$\Phi(\alpha, \beta) = \varphi(f,g)\frac{D(f,g)}{D(\alpha,\beta)}$$

la probabilité s'exprimera par la formule

$$\iint_{\Sigma} \Phi(\alpha,\beta)\,d\alpha\,d\beta$$

en désignant par Σ l'aire (supposée ne pas se recouvrir elle-même) qui correspond dans le champ des variables α, β à l'aire S dans le champ des variables x, y.

On peut, d'une infinité de manières (sous des conditions de continuité pour φ que nous n'énoncerons pas), déterminer le changement de variables de manière que l'on ait

$$\Phi(\alpha,\beta) = 1.$$

La probabilité devient alors

$$\int_{\Sigma} d\alpha\,d\beta$$

c'est-à-dire est proportionnelle à l'aire Σ. Ces variables α, β ainsi déterminées seront dites variables normales pour la fonction $\varphi(x, y)$ qu'on s'est donnée à priori. Ainsi, pour toute fonction

arbitraire, il y a des variables convenablement choisies qui rendent cette fonction arbitraire égale à l'unité et qui, par suite, peuvent justifier par leur emploi le choix en apparence arbitraire. Mais il ne faut pas perdre de vue que l'arbitraire dans le choix des variables n'a pas plus de base concrète que l'arbitraire dans le choix de la fonction φ x, y ; à un problème déterminé correspondent des variables bien déterminées, qui ne dépendent que de la manière dont est posé le problème.

48. Cas où le résultat est indépendant du choix de la fonction arbitraire. — Quel que soit l'intérêt analytique que peuvent avoir les considérations précédentes, il ne saurait suffire à justifier l'étude systématique de l'introduction des fonctions arbitraires s'il ne se produisait fréquemment une circonstance très remarquable, sur laquelle M. Poincaré a le premier attiré l'attention : dans certains cas, le résultat final du calcul est, dans une large mesure, *indépendant du choix de la fonction arbitraire* ; il suffit de supposer que cette fonction satisfait à des conditions assez larges relatives à sa continuité ou au sens de sa variation. Nous allons nous rendre compte de ce fait important par l'étude de quelques problèmes.

Problème XXXII. *Un cadran circulaire est divisé en $2n$ parties égales, peintes alternativement en blanc ou en noir ; une aiguille est mobile autour du centre du cadran et est lancée avec une force suffisante pour faire, avant de s'arrêter, plusieurs fois le tour du cadran : quelle est la probabilité pour qu'elle s'arrête en face d'une division blanche du cadran ?*

On remarquera que l'énoncé ne mentionne pas la position initiale de l'aiguille ; le résultat est en effet, sensiblement indépendant de cette position grâce à l'hypothèse sur la force avec laquelle est lancée l'aiguille ; si celle-ci avait un élan très faible, ne lui faisant parcourir que 2 ou 3 divisions, il n'en serait pas de même.

Désignons par θ l'angle total dont tourne l'aiguille : cet angle augmente de 2π à chaque tour complet du cadran ; d'après l'hypothèse, il dépasse toujours les premiers multiples entiers de

2π. Supposons que ce soit toujours une même personne qui lance l'aiguille ; nous pourrons désigner par $\varphi(\theta)\,d\theta$ la probabilité pour que l'angle total dont elle tourne soit compris entre θ et $\theta + d\theta$: la fonction $\varphi(\theta)$ est nulle lorsque θ est inférieur à 2π, puisque l'aiguille fait toujours plus d'un tour ; elle sera nulle aussi lorsque θ dépassera $100\,\pi$, si l'on suppose que la personne qui lance l'aiguille, en employant toute sa force, est incapable de faire faire à l'aiguille plus de 50 tours ; lorsque θ sera compris entre 2π et 100π, nous ne ferons sur la fonction $\varphi(\theta)$ que des hypothèses de *continuité* ; ces hypothèses sont naturelles si si l'on suppose l'appareil bien construit au point de vue de l'égalité des résistances passives qui ralentissent le mouvement de l'aiguille ; si, par suite d'une défectuosité de l'appareil, une résistance plus grande se produisait dans le voisinage d'une certaine valeur de θ, les arrêts seraient plus fréquents dans cette région et, si cette région correspondait à une division blanche du cadran, la probabilité pour l'arrêt en face de cette division se trouverait accrue.

Notre hypothèse sur θ revient donc à ceci ; la fonction

$$y = \varphi(x)$$

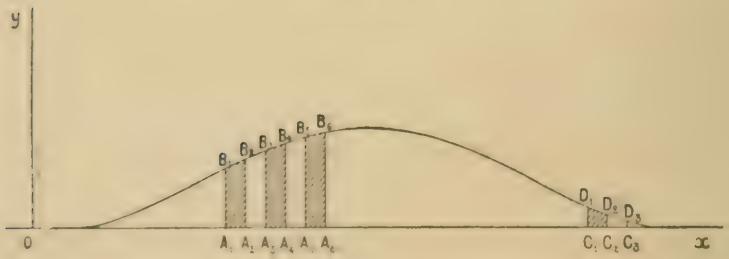

Fig. 17.

peut être représentée par une courbe continue telle que celle que nous figurons (*fig.* 17) ; on a d'ailleurs

$$\int_{-\infty}^{+\infty} \varphi(x)\,dx = 1.$$

Nous avons figuré cette courbe avec un seul maximum ; le raisonnement serait analogue s'il y en avait plusieurs, en nombre limité.

Prenons des abscisses correspondant au début et à la fin de chaque division du cadran à partir de l'instant initial et menons les ordonnées correspondantes : nous figurons seulement quelques-unes de ces ordonnées A_1B_1, A_2B_2, ... A_nB_n. Par hypothèse l'intervalle compris entre A_1B_1 et A_2B_2 correspondra à une division noire du cadran : nous le couvrons de hachures ; l'intervalle compris entre A_2B_2 et A_3A_3 correspond à une division blanche, etc.

Désignons par p la probabilité pour que l'aiguille s'arrête en face d'une division blanche et par q la probabilité pour qu'elle s'arrête en face d'une division noire. On aura évidemment

$$\frac{p}{q} = \frac{\Sigma \text{ aire } A_2B_2A_3B_3}{\Sigma \text{ aire } A_1B_1A_2B_2}$$

le signe Σ du numérateur correspondant à toutes les aires non couvertes de hachures et celui du dénominateur à toutes les aires couvertes de hachures.

Or, il est facile de prouver que le rapport $\frac{p}{q}$ est très voisin de 1, d'autant plus voisin de 1 que n est plus grand. On peut le voir par plusieurs raisonnements.

Tout d'abord, basons-nous seulement sur la continuité : le rapport de deux aires voisines est très voisin de 1, car la fonction $\varphi(x)$ varie très peu lorsque x varie de $\frac{2\pi}{n}$, si l'on suppose n très grand et cette fonction continue. Il est vrai qu'aux extrémités où la courbe se rapproche de Ox une petite différence absolue peut amener un rapport différant sensiblement de 1 : voir sur la figure les aires $C_1D_1C_2D_2$ et $C_2D_2C_3D_3$: mais ces aires très petites peuvent être négligées sans erreur sensible : le rapport $\frac{p}{q}$ se présente alors sous la forme

$$\frac{p}{q} = \frac{\Sigma p_n}{\Sigma q_n}$$

chacun des rapports $\frac{p_n}{q_n}$ différant très peu de l'unité : le rapport $\frac{p}{q}$ diffère donc lui même très peu de l'unité.

Un raisonnement peut-être plus satisfaisant fait intervenir le sens de variation de $\varphi(x)$; sur notre figure, on voit que l'on a :

$$\text{aire } A_1B_1A_2B_2 < \text{aire } A_2B_2A_3B_3 < \text{aire } A_3B_3A_4B_4 < \dots$$

Si donc on désigne les diverses aires correspondant à des secteurs blancs par p_1, p_2, ... p_n, ... et les aires correspondant à des secteurs noirs par q_1, q_2, ..., q_n, ... on aura en supposant que la première aire à gauche soit blanche

$$p_1 < q_1 < p_2 < q_2 < \dots < p_n < q_n < \dots$$

ces inégalités ayant lieu jusqu'au maximum de la courbe ; ensuite, elles changeront de sens. On conclut de là

$$\frac{p_1}{q_1} < 1 ; \quad \frac{p_2}{q_2} < 1 ; \quad \frac{p_3}{q_3} < 1 ; \dots ; \frac{p_n}{q_n} < 1$$

et par suite

$$\rho_n = \frac{p_1 + p_2 + \dots + p_n}{q_1 + q_2 + \dots + q_n} < 1.$$

Mais, d'autre part, on a

$$\frac{p_2}{q_1} > 1 ; \quad \frac{p_3}{q_2} > 1 \dots ; \frac{p_{n+1}}{q_n} > 1$$

et par suite

$$\rho_n' = \frac{p_2 + p_3 + \dots + p_{n+1}}{q_1 + q_2 + \dots + q_n} > 1.$$

Calculons la différence entre les rapports ρ_n' et ρ_n ; on a

$$\rho_n' - \rho_n = \frac{p_{n+1} - p_1}{q_1 + q_2 + \dots + q_n}$$

cette différence est d'autant plus petite que n est plus grand ; nous la désignerons par ε_n ; or nous avons

$$\rho_n < 1 < \rho'_n$$
$$\rho'_n - \rho_n = \varepsilon_n.$$

On en conclut que chacune des différences $\rho'_n - 1$ et $1 - \rho_n$ est inférieure à ε_n. Le même raisonnement s'applique aux rap-

ports analogues relatifs à la partie décroissante de la courbe et par suite au rapport total. La probabilité demandée dans l'énoncé est donc $\frac{1}{2}$; ce résultat est approximatif, avec les hypothèses faites, mais il est, dans une large mesure, indépendant du choix de la fonction arbitraire.

Problème XXXIII. *On établit une table numérique des valeurs d'une fonction quelconque* $\varphi(x)$ *pour des valeurs de* x *croissant par intervalles égaux très petits (par exemple pour* $x = 0$; 0,00001 ; 0,00002 ; 0,00003, etc). *Quelle est la probabilité pour que la deuxième décimale des nombres inscrits dans la table soit le chiffre* 7 ? Cet énoncé a ceci de particulier qu'il renferme lui-même une fonction arbitraire $\varphi(x)$: il ne sera pas nécessaire d'en introduire une. On utilisera le théorème des accroissements finis :

$$\varphi(x_2) - \varphi(x_1) = (x_2 - x_1)\,\varphi'(\xi)$$

dans lequel ξ désigne un nombre compris entre x_1 et x_2. Pour que la deuxième décimale de $\varphi(x)$ soit un 7, il faut par exemple que φx soit compris [1] entre 1,37 et 1,38 ; l'intervalle de variation correspondant pour x est, d'après la formule précédente

$$x_2 - x_1 = \frac{0,01}{\varphi'(\xi)}.$$

Si l'on suppose que x varie par intervalles égaux à 0,000 01, le nombre des valeurs de x comprises dans l'intervalle précédent sera égal à la partie entière du quotient [2]

$$\frac{1\,000}{\varphi'(\xi)}.$$

---

(1) Si l'on a une table à 4 décimales, par exemple, et que l'on ait, suivant l'usage, *forcé* la quatrième décimale dans les cas où la partie négligée dépassait un demi dix-millième, le chiffre des centièmes est un 7 lorsque la vraie valeur de $\varphi(x)$ est comprise, par exemple, entre 1,36995 et 1,37995 ; cela ne change rien aux conclusions.

(2) Nous omettons une discussion arithmétique qui conduirait, dans certains cas, à prendre cette partie entière *diminuée d'une unité.*

Le nombre ξ est tel que $\varphi(\xi)$ est compris entre 1,37 et 1,38 : désignons-le par ξ_7 ; le nombre x_7 des valeurs de x qui donneront un deuxième chiffre décimal égal à 7 sera, d'après cela, sensiblement égal à

$$x_7 = \frac{1\,000}{\varphi(\xi_{17})} + \frac{1\,000}{\varphi(\xi_{27})} + \frac{1\,000}{\varphi(\xi_{37})} + \dots$$

le nombre des termes étant déterminé par l'étendue des variations de la fonction φ dans le champ de la table. De même le nombre n_8 des valeurs de x pour lesquelles le deuxième chiffre décimal est 8 sera sensiblement

$$x_8 = \frac{1\,000}{\varphi(\xi_{18})} + \frac{1\,000}{\varphi(\xi_{28})} + \dots$$

Le rapport entre la probabilité p_7 pour que le chiffre soit 7 et la probabilité p_8 pour que le chiffre soit 8 est égale au rapport $\frac{x_7}{x_8}$; or, si l'on pose

$$\frac{1}{\varphi(x)} = \Phi(x)$$

on a :

$$\frac{x_7}{x_8} = \frac{\Sigma\Phi(\xi_{37})}{\Sigma\Phi(\xi_{38})}$$

et l'on prouvera, *en supposant la continuité* de $\Phi(x)$, que ce rapport est très voisin de *un*, par un raisonnement analogue à celui du problème précédent, et sur lequel nous n'insisterons pas. La probabilité p_7 est donc *un dixième*, *approximativement*.

Comme application on peut, dans une table de logarithmes à sept décimales, chercher combien de fois un chiffre déterminé revient à la cinquième décimale, dans une portion assez étendue de la table : on constatera que la probabilité de rencontrer un chiffre déterminé est bien égale à un dixième.

49. Application à la distribution des petites planètes. — On se propose de déterminer comment les petites planètes se distribuent sur le zodiaque : en simplifiant un peu les termes du problème, on peut le ramener au suivant.

Problème XXXIV. *On considère n points qui se déplacent sur un même cercle d'un mouvement uniforme, mais avec des vitesses différentes : quelle est l'espérance mathématique d'une personne qui doit toucher autant de francs qu'il y a de points sur un arc donné ?* On remarquera qu'à un instant donné, la probabilité pour un point de se trouver sur un arc donné est proportionnelle à la longueur α de cet arc, car puisque le mouvement de chaque point est uniforme, le temps pendant lequel il se trouve sur chaque arc est proportionnel à la longueur de cet arc ; l'espérance mathématique est donc $\frac{\alpha}{2\pi}$ pour un point : elle est $\frac{n\alpha}{2\pi}$ si l'on considère n points. On pourrait se contenter de cette solution que j'ai discutée dans un mémoire auquel je me contente de renvoyer [1] ; examinons celle que propose M. Poincaré : il recherche la valeur moyenne d'une fonction périodique de l'abscisse curviligne α de l'un des points, disons par exemple de $\sin\alpha$; il est clair que cette valeur moyenne sera d'autant plus voisine de zéro que la distribution des points sera plus régulière. Or, si l'on désigne par a_k l'abscisse de l'un des points à l'origine des temps et par b_k sa vitesse angulaire, son abscisse curviligne est $a_k + b_k t$ et la valeur moyenne que l'on cherche est

$$\frac{1}{n}\,\Sigma \sin (a_k + b_k t).$$

La méthode de M. Poincaré consiste, les points étant très nombreux, à remplacer cette somme par une intégrale : désignons par $\varphi(x, y)\,dx\,dy$ le nombre de ces points pour lesquels a_k est compris entre x et $x + dx$ et b_k compris entre y et $y + dy$: on pourra écrire

$$\Sigma \sin (a_k + b_k t) = \iint \sin (x + ty)\,\varphi(x, y)\,dx\,dy$$

$$n = \iint \varphi(x, y)\,dx\,dy$$

(1) *Sur les principes de la théorie cinétique des gaz.* (Annales de l'École Normale, 1906). La figure 18 ci-après, dont le cliché m'a été obligeamment prêté par M. Gauthier-Villars, est empruntée à ce Mémoire.

les intégrales doubles étant étendues à tout le champ dans lequel $\varphi(x, y)$ n'est pas nul ; la valeur moyenne cherchée a donc pour valeur

$$\frac{\iint \sin(x + ty)\, \varphi(x, y)\, dxdy}{\iint \varphi(x, y)\, dxdy}$$

et on démontre facilement que, *lorsque t augmente indéfiniment*, elle tend vers zéro : donc la distribution, si elle n'est pas régulière à l'origine, tend à le devenir.

J'ai tenu à indiquer ce raisonnement, qui me paraît toutefois pouvoir être soumis à quelques objections que j'ai développées dans le mémoire cité plus haut. Mais ces objections étant tout à fait analogues à celles qui vont se présenter dans les problèmes qui se rattachent à la théorie cinétique des gaz, je n'y insisterai pas.

50. Application à la théorie cinétique des gaz. — L'étude du mouvement de n molécules dans l'espace à 3 dimensions peut être ramenée à l'étude du mouvement d'un point unique dans l'espace à $3n$ dimensions : ce problème fort général peut être éclairé par l'étude du mouvement d'un point dans un espace à deux dimensions seulement, c'est-à-dire dans le plan ; nous nous bornerons ici à cette étude, renvoyant pour le surplus au Mémoire déjà cité.

Considérons donc un point matériel assujetti à se déplacer dans un plan d'un mouvement rectiligne et uniforme ; soient x et y ses coordonnées rectangulaires ; nous allons chercher quelle est la probabilité pour que la partie entière de chacun des nombres x et y soit un nombre pair. En d'autres termes, chacun de ces nombres doit être compris entre 0 et 1, ou entre 2 et 3, ou entre 4 et 5, etc.

Si nous divisons en carrés de côtés égaux à l'unité de longueur la partie du plan qui correspond aux coordonnées positives, nous pourrons couvrir de hachures les carrés pour lesquels la condition précédente est remplie et dire simplement que nous

cherchons la probabilité pour que le point variable se trouve à l'intérieur d'un carré couvert de hachures (*fig.* 18).

Il n'est pas inutile d'observer que si l'on considère un point matériel assujetti à se mouvoir à l'intérieur d'un carré d'un mouvement rectiligne et uniforme, en se réfléchissant sur les parois suivant la loi classique, son mouvement est entièrement analogue à celui du point que nous considérons : pour rendre identiques les deux problèmes, il suffit de supposer que l'on observe le mouvement du point intérieur au carré dans des miroirs plans parfaitement polis passant par les côtés du carré et perpendiculaires à son plan. On se rendra compte aussi que le problème est le même si nous considérons un point matériel intérieur à un cube et se réfléchissant sur ses parois suivant la loi classique, et que nous portions notre attention sur la projection de ce point matériel sur une face du cube. Mais nous nous bornerons à l'énoncé donné en premier lieu.

Soient α et β les coordonnées du point considéré à l'origine du temps, u et v les projections sur les axes de sa vitesse ; ses coordonnées à l'époque t sont données par les formules

$$x = \alpha + ut$$
$$y = \beta + vt.$$

Je laisserai de côté la discussion, fort intéressante à certains égards, de ces formules lorsque l'on suppose que α, β, u, v sont des nombres rigoureusement connus : les propriétés arithmétiques de ces nombres (le fait qu'ils sont commensurables ou incommensurables) interviennent dans cette discussion. Ce n'est pas ainsi que peut se poser un problème pratique, car supposer un nombre connu avec une précision indéfinie est une conception purement théorique, qui ne peut avoir aucune signification réelle. En fait, on doit regarder α, β, u, v comme connus seulement avec une certaine approximation ; pour abréger les écritures, nous supposerons α et β exactement connus, car cette hypothèse ne change rien au résultat : nous désignerons par ε et η les erreurs en plus ou en moins commises sur u et v, de sorte que les formules précédentes doivent être remplacées par les suivantes :

$$x = \alpha + u't \quad |u' - u| \leqslant \varepsilon$$
$$y = \beta + v't \quad |v' - v| \leqslant \eta$$

Soit ω le point de coordonnées α, β et ξ le point de coordonnées $\alpha + u$, $\beta + v$. Lorsque u et v prennent toutes les valeurs possibles, c'est-à-dire lorsque u varie de $u - \varepsilon$ à $u + \varepsilon$ et v de $v - \eta$ à $v + \eta$, le point ξ se déplace à l'intérieur d'un rectangle MNPQ dont les côtés ont respectivement pour longueurs 2ε et 2η. Par hypothèse le point ξ peut occuper une position quelconque à l'intérieur de ce rectangle. Nous supposerons que la probabilité pour qu'il se trouve dans une certaine aire est proportionnelle à cette aire : on pourrait faire voir, en appliquant la méthode de M. Poincaré, exposée au n° **48** que cette hypothèse n'est pas essentielle ; mais elle simplifiera les raisonnements et les calculs.

Il est maintenant aisé de voir où se trouve le point ξ' ayant pour coordonnées

$$x = \alpha + u't$$
$$y = \beta + v't$$

ce point est homothétique de ξ, le centre d'homothétie étant le point ω et le rapport d'homothétie étant égal à t ; il se trouve donc

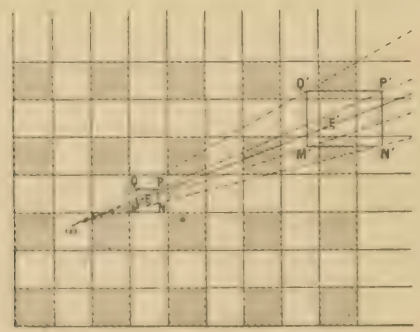

Fig. 18.

à l'intérieur d'un rectangle M'N'P'Q', homothétique du rectangle MNPQ dans le même rapport. Nous avons dû faire la figure en supposant t égal seulement à quelques unités ; mais il est aisé de se rendre compte de ce qui se passerait si l'on supposait t égal à quelques milliers ou à quelques millions de fois l'unité. Quelque petit que soit le rectangle MNPQ, le rectangle M'N'P'Q'

dont les dimensions sont t fois plus grandes, finira par être extrêmement grand par rapport au carrelage, et dès lors les carrés couverts de hachures occuperont à peu près le quart de sa surface totale. Ce résultat sera d'autant plus exact que le rectangle sera plus grand, car les surfaces des carrés partiellement intérieurs et partiellement extérieurs deviendront négligeables par rapport aux aires des carrés entièrement intérieurs.

On voit quel est le mécanisme de l'action du temps dans l'égalisation des probabilités, lorsqu'il y a dans les données une indétermination, si petite soit-elle : pour de petites valeurs de t, par exemple dans le cas du rectangle $MNPQ$ représenté sur la figure, il pourra arriver que la portion couverte de hachures soit inférieure au quart de la surface totale : la probabilité pratique n'est pas égale à la probabilité théorique ; mais il y a tendance à l'égalisation entre ces probabilités lorsque t prend des valeurs de plus en plus grandes. C'est en vertu d'un mécanisme analogue que les vitesses des molécules d'un gaz, telles qu'on les considère dans la théorie cinétique, tendent vers la répartition que le calcul démontre devoir être la plus probable. On ne peut pas démontrer que cette répartition la plus probable est effectivement réalisée à un instant donné et, si elle *n'est pas* réalisée, on doit dire que sa probabilité pratique en cet instant est *zéro* ; mais on peut démontrer que, quelle que soit la répartition à une époque donnée, la probabilité pour qu'à une époque postérieure la répartition que le calcul indique comme probable soit effectivement réalisée est tellement voisine de l'unité qu'on peut la confondre pratiquement avec la certitude. Le même raisonnement s'appliquerait aux époques antérieures, avec toutefois la différence suivante : lorsque l'on se propose de calculer l'état futur en partant de l'état actuel, on doit admettre que l'état actuel n'est pas rigoureusement connu, et de très petites variations, intérieures aux limites entre lesquelles l'état actuel est inconnu, suffisent pour la définition et le calcul des probabilités diverses pour l'état futur : car on ne peut parler rigoureusement de probabilités que s'il y a une part d'inconnu. Au contraire, lorsque l'on se propose de remonter de l'état actuel à l'état antérieur, l'ignorance est moins grande dans bien des cas, dans les cas précisément où ce que l'on sait de l'état anté-

rieur nous renseigne sur l'état actuel avec plus de précision que ne saurait nous en donner une étude directe de cet état actuel. Par exemple si l'on a deux récipients dont l'un était vide et l'autre plein de gaz et qu'on les ait mis en communication en ouvrant un robinet, l'état actuel ne permettrait de remonter à l'état antérieur au moment de l'ouverture du robinet que s'il était connu avec une précision dépassant de beaucoup les possibilités expérimentales, car la plus minime modification conduira, on s'en convainc aisément, à un état antérieur tout différent ; il n'est donc pas possible d'employer ici pour définir et calculer la probabilité de cet état antérieur, une méthode analogue à celle que j'ai esquissée tout à l'heure.

CHAPITRE IX

ERREURS D'OBSERVATION. LOI DE GAUSS

51. Délimitation du problème. — Nous ne traiterons pas toutes les questions soulevées par une théorie complète des erreurs d'observation, théorie qui exigerait à elle seule un livre entier. En particulier, nous ne donnerons aucun détail sur les méthodes de calcul numérique qu'il convient d'employer pour déduire d'un certain nombre d'observations les valeurs des quantités que ces observations avaient pour but de déterminer. Nous n'insisterons pas non plus sur les discussions théoriques dont a été l'objet la loi de Gauss et sur les développements mathématiques auxquels ces discussions ont donné lieu ; il ne semble pas en effet que les résultats obtenus dans cette voie aient une importance en rapport avec l'effort analytique qu'ils exigent ; et il semble que ce sont les raisons les plus simples qui rendent le mieux compte de la quasi-universalité pratique de la loi de Gauss (1).

(1) La raison principale pour laquelle il convient, à mon avis, de n'employer qu'avec beaucoup de prudence en ces questions les raisonnements analytiques compliqués (intégration d'équations différentielles, d'équations fonctionnelles, etc.) est la suivante : il ne peut être question que de raisonnements approchés, valables seulement sous certaines conditions qui ne sont jamais *toutes* remplies; dès lors, chaque équation comporte une erreur, souvent très faible en pratique, mais jamais nulle. Or si l'on peut limiter assez facilement les erreurs sur les équations qui se déduisent algébriquement d'équations comportant elles-mêmes des erreurs, il n'en est pas de même pour les équations différentielles et les équations fonctionnelles : une modification très faible dans une telle équation peut changer du tout au tout la nature analytique de l'intégrale.

Nous nous bornerons donc à indiquer ces raisons et nous développerons ensuite les principales conséquences théoriques de la loi de Gauss. Ainsi des trois parties que comporterait une théorie complète : établissement de la loi de Gauss, son étude théorique et ses applications pratiques, nous laisserons de côté entièrement la troisième et en grande partie la première, nous bornant à la seconde.

52. Enoncé de la loi de Gauss. — On cherche à mesurer une quantité, par exemple une longueur, dont la vraie valeur est a : si la mesure donne une valeur a', l'erreur x est $a' - a$; la loi de Gauss consiste en ce que la probabilité pour que cette erreur soit comprise entre x et $x + dx$ est égale à

$$\frac{k}{\sqrt{\pi}} e^{-k^2x^2} dx,$$

k étant une constante que l'on appelle *précision* du mode de mesure employé.

L'énoncé même de cette loi appelle plusieurs observations.

Tout d'abord, qu'entendra-t-on par la vraie valeur de la quantité à mesurer ? Il arrivera souvent que cette valeur n'est connue que précisément à la suite d'une série de mesures : n'y a-t-il pas un cercle vicieux à évaluer l'erreur de ces mesures à l'aide d'un nombre fourni par ces mesures elles-mêmes. Nous n'insisterons pas ici sur cette difficulté qui peut être rattachée à la théorie de la probabilité des causes : indiquons seulement qu'*il y a des cas où cette difficulté ne se présente pas* : on peut admettre que l'on se borne, au moins provisoirement, à ces cas : ce sont ceux où la quantité à mesurer est connue, par des mesures antérieures, combinées ou non avec des déductions théoriques, *avec une approximation beaucoup plus grande que celle des mesures actuelles*. C'est ce qui se produit si l'on fait effectuer une manipulation de physique ou de chimie dans laquelle de nombreux élèves doivent déterminer une constante par ailleurs bien connue.

On peut en dire autant de la constante k, qui figure dans la formule de Gauss ; c'est seulement à la suite des observations

qu'elle est déterminée ; nous indiquerons comment peut se faire cette détermination.

La loi de Gauss donne une égale probabilité aux erreurs positives et aux erreurs négatives, lorsque leur valeur absolue est la même ; c'est seulement dans les cas où cette symétrie est réalisée que l'on peut supposer qu'elle sera vérifiée.

Observons enfin que si l'on remplace x par une fonction de x la forme analytique de la loi se trouve modifiée ; c'est là une objection analogue à celles que nous avons déjà rencontrées ; on doit y répondre de la même manière, par l'observation des faits : si l'on se sert d'une règle divisée pour évaluer une longueur, les diverses parties de cette règle sont homogènes, et ce n'est pas le carré de la longueur que l'on mesure, mais la longueur elle-même ; c'est donc sur la longueur et non sur son carré que portent les erreurs.

D'ailleurs, lorsque les erreurs sont petites, la formule de Gauss conserve sa forme *à la première approximation* : posons

$$b = \varphi(a) \qquad b' = \varphi(a')$$
$$x = a' - a \qquad y = b' - b$$

Nous aurons :

$$y = \varphi'(\alpha)\, x$$

en désignant par α un nombre compris entre a et a' : si toutes les mesures a sont comprises dans un intervalle assez petit, on a sensiblement $\varphi'(\alpha) = \varphi(a) = c$, et il vient

$$\frac{k}{\sqrt{\pi}} e^{-k^2x^2} dx = \frac{k}{c\sqrt{\pi}} e^{-\frac{k^2}{c^2}y^2} dy$$

La précision est simplement divisée par c. Supposons par exemple que l'on mesure une longueur de 100 mètres et que les diverses mesures diffèrent entre elles de moins de 1 mètre : si au lieu de la mesure, on introduit son carré, on a

$$\varphi(a) = a^2 \qquad \varphi'(a) = 2a$$

et l'on voit que la valeur de c est comprise entre 199 et 201 : on commet une erreur relative inférieure à $\frac{1}{200}$ en prenant $c = 200$.

53. Fondement de la loi de Gauss. — On pourrait se borner à dire que la loi de Gauss se justifie dans certains cas par l'expérience ; on peut donc convenir d'appeler *normales* les séries de mesures auxquelles la loi de Gauss s'applique, et *anormales* les séries auxquelles cette loi ne s'applique pas. Ces définitions posées, on fera la théorie des séries de mesures normales, et on sera parfaitement justifié à la fonder sur la loi de Gauss. Certaines personnes estimeront qu'en procédant ainsi, on a simplement fait usage du droit — elles diront même du devoir — qu'a tout savant de définir pour son usage un langage clair et commode; d'autres, plus sévères, penseront que l'on a simplement déplacé la difficulté, que ne saurait résoudre une convention de langage, si ingénieuse fût-elle, et qu'il s'agit précisément de savoir pourquoi il arrive assez fréquemment que des séries de mesures sont normales : c'est ce que nous allons essayer d'expliquer brièvement ; nous laisserons de côté une autre question plus difficile, sur laquelle nous reviendrons au Livre III (voir n° **72**), à savoir : quel est le caractère intrinsèque des séries normales ? C'est seulement, en effet, lorsqu'on aura trouvé aux séries normales un autre caractère que celui de satisfaire à la loi de Gauss par lequel on les a définies, que cette définition sera réellement instructive et que sa portée dépassera l'utilité d'un langage abrégé.

Le fait que certaines séries de mesures sont normales s'explique par les résultats que nous avons obtenus au Livre I relativement aux écarts. Nous avons montré en effet que, si l'on fait n_1 tirages dans une urne avec une probabilité p_1, n_2 tirages dans une seconde urne avec une probabilité d'écart p_2, etc., l'écart relatif λ suit une loi normale (loi de Gauss avec $k = 1$), l'écart absolu λU étant égal au produit de l'écart relatif par l'unité d'écart U définie par la relation

$$U^2 = 2n_1p_1q_1 + 2n_2p_2q_2 + \dots$$

Or, que se passe-t-il lorsqu'on effectue une mesure ? Les causes d'erreur sont extrêmement nombreuses ; chacune d'elles a une certaine probabilité et peut produire une erreur déterminée. Lorsqu'une même cause peut produire deux erreurs de valeurs différentes, on peut, par la pensée, regarder ces deux erreurs comme ayant deux causes distinctes, dont les probabi-

lités sont distinctes ; lorsque l'erreur produite par une même cause varie d'une manière continue, on décomposera son champ de variation en un certain nombre d'éléments assez petits pour que l'erreur soit regardée comme constante dans chacun d'eux. Cela étant, on désignera par n_1 le nombre d'erreurs possibles dont la valeur est α_1 et la probabilité p_1, par n_2 le nombre d'erreurs possibles dont la valeur est α_1 et la probabilité p_2, etc. Les nombres α_1, α_2, ... sont positifs ou négatifs et l'on a :

$$n_1 p_1 \alpha_1 + n_2 p_2 \alpha_2 + \ldots = 0,$$

c'est-à-dire que l'erreur totale est nulle si chaque erreur partielle se produit suivant sa probabilité : c'est une conséquence de l'hypothèse, indispensable pour que la loi de Gauss puisse être vraie, que les erreurs négatives sont aussi probables que les erreurs positives.

En réalité les erreurs α_1 se produisent non pas $n_1 p_1$ fois, mais $n_1 p_1 + \lambda_1 u_1$ en désignant par u_1 l'unité d'écart partiel [5] et λ_1 l'unité d'écart, qui satisfait à la loi de Gauss, n° **28**) etc. On en conclut que l'erreur est :

$$\lambda U = \lambda_1 u_1 \alpha_1 + \lambda_2 u_2 \alpha_2 + \ldots + \lambda_k u_k \alpha_k$$

et un raisonnement analogue à celui des n^os^ **33** et **34**, nous donne :

$$U^2 = u^2{}_1 \alpha^2{}_1 + u^2{}_2 \alpha^2{}_2 + \ldots + u^2{}_k \alpha^2{}_k.$$

La quantité U ainsi définie est l'unité d'écart et l'écart relatif λ vérifie la loi de Gauss. On voit que l'unité d'écart U est l'inverse du nombre k que nous avons appelé la précision ; car la probabilité pour λ étant

$$\frac{1}{\sqrt{\pi}} e^{-\lambda^2} d\lambda$$

la probabilité pour $x = \lambda U$ est

$$\frac{1}{U\sqrt{\pi}} e^{-\frac{x^2}{U^2}} dx.$$

---

(5) On a : $u^2{}_1 = m_1 p_1 q_1$.

On a donc

$$\frac{1}{k^2} = \Sigma 2 n_i p_i q_i \varepsilon_i^2,$$

de sorte que la précision est, en gros, inversement proportionnelle à la valeur des erreurs élémentaires et à la racine carrée de leur nombre. Si les erreurs élémentaires deviennent toutes 10 fois plus petites, la précision devient 10 fois plus grande; si les erreurs élémentaires restant du même ordre de grandeur, deviennent 100 fois plus nombreuses, la précision devient 10 fois plus faible. Ces énoncés prendront leur sens complet lorsque nous aurons étudié les relations avec la réalité de la valeur du nombre k, que nous avons appelé précision, ce qui n'est jusqu'ici qu'une définition de mots (voir n° **55**).

En résumé, la loi de Gauss se trouve justifiée par le fait que les erreurs possibles étant nombreuses, leurs combinaisons sont soumises aux lois que nous avons obtenues par l'analyse combinatoire dans le cas où l'on étudie un grand nombre d'épreuves répétées.

54. Invariance de la loi de Gauss. — On peut trouver une confirmation *a posteriori* de la loi de Gauss dans une propriété d'invariance que l'on peut énoncer comme il suit :

Lorsque les erreurs commises sur les mesures de plusieurs quantités satisfont à la loi de Gauss, il en est de même de l'erreur commise sur la somme, si l'on prend pour valeur de la somme la somme des valeurs des mesures.

Il suffira de démontrer cette propriété dans le cas de deux quantités pour lesquelles les probabilités d'une erreur comprise entre x et $x + dx$ et d'une erreur comprise entre y et $y + dy$ sont respectivement

$$\frac{k}{\sqrt{\pi}} e^{-k^2x^2} dx$$
$$\frac{k'}{\sqrt{\pi}} e^{-k'^2y^2} dy.$$

Quelle est la probabilité d'une erreur comprise entre z et $z + dz$ pour la somme des deux quantités? Si l'erreur commise

sur la première est x, l'erreur commise sur la seconde doit être comprise entre $z - x$ et $z - x + dz$; la probabilité d'une telle erreur est

$$\frac{k'}{\sqrt{\pi}} e^{-k'^2 z - x^2} dz.$$

On aura la probabilité totale en multipliant cette expression par la probabilité pour x d'être compris entre x et $x + dx$ et intégrant par rapport à x entre les limites $-\infty$ et $+\infty$, ce qui donne une expression de la forme $J dz$, en posant

$$J = \frac{kk'}{\pi} \int_{-\infty}^{+\infty} e^{-k'^2(z-x)^2 - k^2 x^2} dx$$

ce qui, en utilisant les résultats du n° **33**, s'écrit

$$J = \frac{k''}{\sqrt{\pi}} e^{-k''^2 z^2}$$

la précision k'' étant définie par la relation

$$\frac{1}{k''^2} = \frac{1}{k^2} + \frac{1}{k'^2}.$$

55. Précision d'une série de mesures. — Il est temps de mettre en évidence la signification concrète du coefficient k auquel nous avons donné le nom de précision. La probabilité pour qu'une erreur soit comprise entre $k\lambda_1$ et $k\lambda_2$ est

$$\frac{1}{k\sqrt{\pi}} \int_{k\lambda_1}^{k\lambda_2} e^{-k^2 x^2} dx = \frac{1}{\sqrt{\pi}} \int_{\lambda_1}^{\lambda_2} e^{-\lambda^2} d\lambda$$

c'est-à-dire est indépendante de k. Ainsi lorsque la précision d'un système de mesures est 10 fois plus grande que la précision d'un autre système, cela veut dire que la probabilité d'erreurs comprises entre certaines limites pour le premier système est la même que la probabilité d'erreurs comprises entre des limites 10 fois plus grandes pour le second système : par exemple, les erreurs d'un certain nombre de millimètres dans le premier système auront la même probabilité que les erreurs du même nombre de centimètres dans le second système.

L'*erreur probable* est, par définition, la somme des produits obtenus en multipliant chaque erreur possible par sa probabilité ; si on tient compte du signe des erreurs, l'erreur probable est zéro, car les erreurs positives sont aussi probables que les négatives ; si l'on prend l'erreur en valeur absolue, la valeur de l'erreur probable est

$$\frac{2k}{\sqrt{\pi}} \int_0^\infty e^{-k^2 z^2} z\,dz = \frac{1}{k\sqrt{\pi}}.$$

Elle est, comme il est naturel, d'autant plus petite que la précision est plus grande. Lorsque l'on a n observations, dont nous désignerons les erreurs prises en valeurs absolues par $l_1, l_2, \ldots, l_n$, on aura une valeur approchée de l'erreur probable en prenant la moyenne de ces erreurs, ce qui donne la relation

$$\frac{1}{k\sqrt{\pi}} = \frac{e_1 + e_2 + \ldots + e_n}{n}$$

à l'aide de laquelle on peut calculer pratiquement la précision k.

De même, la valeur probable ou moyenne du carré de l'erreur a pour expression

$$\frac{2k}{\sqrt{\pi}} \int_0^\infty e^{-k^2 z^2} z^2\,dz = \frac{1}{2k^2}$$

et pour valeur approchée

$$\frac{1}{2k^2} = \frac{e^2_1 + e^2_2 + \ldots + e_n{}^2}{n}.$$

On peut améliorer ces valeurs approchées en recherchant la valeur de σ pour laquelle le carré de la différence

$$\frac{1}{k\sqrt{\pi}} - \sigma \frac{e_1 + e_2 + \ldots + e_n}{n}$$

a la plus petite valeur probable : en d'autres termes on cherche à rendre minimum l'expression

$$\mathfrak{M}\left(\frac{1}{k\sqrt{\pi}} - \sigma \frac{e_1 + e_2 + \ldots + e_n}{n}\right)^2$$

Or, on a

$$\mathfrak{M}(1) = 1 \quad , \quad \mathfrak{M}(e_1) = \frac{1}{k\sqrt{\pi}} \quad , \quad \mathfrak{M}(e_1 e_2) = \frac{1}{k^2\pi} \quad , \quad \mathfrak{M}(e_1^2) = \frac{1}{2k^2}$$

De sorte que l'expression à calculer s'écrit

$$\frac{1}{k^2\pi} - \frac{2\sigma}{k^2\pi} + \frac{\sigma^2}{\pi^2}\left(\frac{n}{2k^2} + \frac{n(n-1)}{k^2\pi}\right)$$

c'est-à-dire

$$\frac{1}{k^2\pi}\left[1 - 2\sigma + \sigma^2\left(\frac{\pi - 2}{2n} + 1\right)\right].$$

Le minimum correspond (1) à

$$\sigma = \frac{1}{1 + \dfrac{\pi - 2}{2n}}$$

et a pour valeur

$$\frac{1}{2n\,k^2\,n^2}\left(1 - \frac{2}{\pi}\right)\frac{1}{1 + \dfrac{\pi - 2}{2n}}.$$

Ainsi, si l'on prend

$$\frac{1}{k\sqrt{\pi}} = \frac{1}{1 + \dfrac{\pi - 2}{2n}}\,\frac{e_1 + e_2 + \ldots + e_n}{n},$$

On peut espérer faire une erreur probable plus faible sur la valeur de la précision ainsi obtenue. Cette remarque n'a pas d'ailleurs grand intérêt pratique car, si n est grand, les deux valeurs diffèrent très peu et, si n est petit, elles sont toutes deux très incertaines.

56. Poids d'un résultat. Moyennes. — On donne le nom de *poids* d'une mesure au carré k^2 de sa *précision* k. Cette dénomination est justifiée par la solution du problème suivant.

---

(1) Voir BERTRAND, p. 194. Il y a dans Bertrand une faute d'impression à la valeur du facteur qu'il appelle λ et qui est égal à $\sigma\sqrt{\pi}$.

PROBLÈME XXXVII. *Pour mesurer une longueur z on a fait une première série de mesures, dont la précision est k et une seconde série dont la précision est k' ; à supposer que l'on connaisse seulement le résultat m d'une mesure unique de la première série et le résultat m' d'une mesure unique de la seconde série, quelle valeur doit-on adopter pour z?*

On est conduit, pour des raisons évidentes [1], à adopter pour z une expression de la forme

$$z = \frac{hm + h'm'}{h + h'}$$

et l'on dit alors que h et h' sont les *poids* respectifs attribués aux deux mesures.

Si l'on désigne par x et y les erreurs commises sur m et m' l'erreur commise sur z est

$$\frac{hx + h'y}{h + h'}.$$

Nous ne pouvons pas chercher à rendre minimum la valeur probable de cette expression, car cette valeur probable est zéro, les erreurs x et y étant forcément prises avec leurs signes ; nous chercherons à rendre minimum la valeur probable de son carré, c'est-à-dire

$$\mathfrak{M}\left(\frac{hx + h'y}{h + h'}\right)^2 = \frac{h^2\,\mathfrak{M}(x^2) + h'^2\,\mathfrak{M}(y^2)}{(h + h')^2}$$

car la valeur moyenne de xy est nulle. Or, on a :

$$\mathfrak{M}(x^2) = \frac{1}{2k^2}$$

$$\mathfrak{M}(y^2) = \frac{1}{2k'^2}$$

---

(1) En effet, l'expression de z doit satisfaire aux conditions suivantes :

1° Si l'on multiplie m et m' par une même constante, elle est multipliée par cette constante : elle est donc homogène et du premier degré.

2° Si l'on ajoute à m et m' une même constante, z est augmenté de cette constante ; l'expression est donc linéaire en même temps qu'homogène.

3° Si $m = m'$, z doit être égal à leur valeur commune.

et l'on doit chercher le minimum de

$$\left(\frac{h^2}{k^2}+\frac{h'^2}{k'^2}\right)\frac{1}{(h+h')^2}.$$

Si l'on pose

$$h = u\,(h+h')$$
$$h' = (1-u)\,(h+h')$$

cette expression devient

$$\frac{u^2}{k^2}+\frac{(1-u)^2}{k'^2}=0$$

et en annulant sa dérivée par rapport à u, on obtient

$$\frac{u}{k^2}-\frac{1-u}{k'^2}=0$$

c'est-à-dire

$$\frac{h}{k^2}=\frac{h'}{k'^2}.$$

Les *poids* h et h' sont donc proportionnels à k^2 et k'^2 et la valeur à choisir pour z est

$$z=\frac{k^2m+k'^2m'}{k^2+k'^2}$$

Lorsque plusieurs observations appartiennent à la même série et ont par suite la même précision, on est ainsi conduit à prendre leur moyenne arithmétique.

D'une manière générale, soient n_1 observations de précision k_1 et dont le résultat moyen est α_1, n_2 observations de précision k_2 et dont le résultat moyen est α_2, etc., la valeur à adopter pour z sera

$$z=\frac{n_1k_1^2\alpha_1+\ldots+n_pk_p^2\alpha_p}{n_1k_1^2+\ldots+n_pk_p^2}$$

c'est-à-dire que le poids d'un ensemble d'observations s'obtient en multipliant le poids commun de ces observations par leur nombre. On en conclut que *la précision de la moyenne de*

plusieurs observations s'obtient en multipliant la précision d'une observation unique par la racine carrée du nombre des observations. Cet important résultat peut s'établir directement ; si l'on pose :

$$x = x_1 + x_2 + \ldots + x_n$$

et que l'on désigne par k_1, k_2, ... k_n et k les précisions relatives aux erreurs x_1, x_2, ... x_n et à leur somme x, on a

$$\frac{1}{k^2} = \frac{1}{k_1^2} + \frac{1}{k_2^2} + \ldots + \frac{1}{k_n^2}.$$

Or, si l'on suppose $k_1 = k_2 = \ldots = k_n$ cette formule donne

$$k = \frac{k_1}{\sqrt{n}}.$$

Mais la précision sur $\frac{x}{n}$ est n fois plus grande que la précision sur x, car à une erreur commise sur x correspond une erreur n fois plus faible sur $\frac{x}{n}$: la précision dans l'erreur sur la moyenne égale à

$$\frac{x}{n} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

est donc

$$nk = k_1\sqrt{n}.$$

Ainsi, par la moyenne de 100 mesures, on obtient une précision 10 fois plus grande que par une mesure unique, par la moyenne de 10 000 mesures, on obtient une précision 100 fois plus grande que par une mesure unique. Nous laissons de côté la question de savoir quelle est la limite de la précision qui peut ainsi *pratiquement* être atteinte en multipliant les mesures [1].

---

(1) Dans mon cours de 1908-1909, j'ai étudié à ce point de vue diverses questions expérimentales, et en particulier de très intéressantes expériences sur les fils employés à la mesure des bases géodésiques, d'après les publications du Bureau international des poids et mesures et les indications qu'ont bien voulu me donner très aimablement M. Benoît et M. Ch. Ed. Guillaume. La conclusion pratique à laquelle j'ai été conduit, c'est que l'on peut raisonnablement espérer, en combinant quelques centaines de mesures, obtenir une décimale exacte de plus que par un petit nombre de mesures ; mais qu'il est difficile d'aller plus loin, par suite des difficultés pratiques qu'il y a à être assuré qu'on mesure *la même chose* lorsque le nombre des mesures atteint plusieurs milliers.

LIVRE III

PROBABILITÉS DES CAUSES

CHAPITRE X

CAS DES PROBABILITÉS DISCONTINUES

57. Quelques problèmes simples.

Problème XXXVIII. *Deux urnes* A *et* B *renferment, l'urne* A : 3 *boules blanches et* 1 *noire ; l'urne* B, 1 *boule blanche et* 3 *noires. On extrait une boule d'une des urnes, choisie au hasard et on constate qu'elle est blanche ; quelle est la probabilité pour que l'urne choisie soit* A ?

Chacune des urnes renfermant le même nombre de boules, la probabilité d'extraire une quelconque des boules est la même : comme il y a 3 boules blanches dans A et 1 dans B, la probabilité pour que la boule blanche extraite soit l'une des boules de A est $\frac{3}{4}$; telle est la probabilité demandée.

On peut aussi raisonner comme il suit : imaginons 40 systèmes de deux urnes $A_1, A_2, \dots A_{40}$; $B_1, B_2, \dots B_{40}$ couplées deux à deux, toutes les urnes A, ont la même composition que l'urne A et toutes les urnes B_i ont la même composition que l'urne B. Supposons que l'on extraie une boule au hasard d'une urne de chaque couple choisie au hasard ; on peut présumer que l'urne choisie sera A dans 20 des couples et B dans les 20 autres ; or, si l'on extraie 20 fois une boule d'une urne dont la composition est A, l'hypothèse la plus probable est que l'on obtiendra 15 blanches et 5 noires ; de même on obtiendra 5 blanches et

15 noires dans les 20 extractions d'urnes B; ce qui permet de former le tableau suivant

20 extr. A	donnent	15 blanches	5 noires
20 » B	»	5 »	15 »
40 extr.	»	20 »	20 »

On voit que les 20 extractions ayant fourni les blanches se décomposent en 15 provenant d'urnes A et 5 provenant d'urnes B; la probabilité pour que l'une d'elles provienne d'une urne A est donc $\frac{3}{4}$.

Problème XXXIX. *Les urnes* A *et* B *étant composées comme dans le problème précédent, on fait deux extractions de chaque urne, en ayant soin de remettre la boule extraite après la première extraction pour ne pas changer la composition de l'urne; sachant que chacune des deux extractions donne une boule blanche, quelle est la probabilité pour que l'urne choisie soit A.*

La probabilité d'extraire deux boules blanches de A est $\frac{9}{16}$; la probabilité de les extraire de B est $\frac{1}{16}$; si donc on fait 32 expériences, dans 16 desquelles on choisit A et dans 16 desquelles on choisit B, les nombres probables de cas où l'on extraira 2 blanches seront respectivement 9 et 1; il y a donc 9 chances sur 10, lorsque cette éventualité se produit pour que l'on ait choisi l'urne A.

Plus généralement, si n tirages consécutifs donnaient une blanche, on pourrait parier 3^n contre 1 que l'urne choisie est A.

Problème XL. *On a trois coffrets dont chacun a deux tiroirs; le premier coffret renferme dans chaque tiroir une pièce d'or, le second renferme une pièce d'argent et une pièce d'or et le troisième deux pièces d'argent. On ouvre l'un des tiroirs et on y trouve une pièce d'or; quelle est la probabilité pour que le second tiroir du même coffret renferme une pièce d'or?*

La question revient à la suivante : quelle est la probabilité pour que le tiroir ouvert appartienne au premier coffret? Or,

trois tiroirs renfermaient une pièce d'or ; la probabilité pour que l'on ait ouvert chacun d'eux est donc $\frac{1}{3}$, car ces 3 tiroirs sont supposés extérieurement identiques ; la probabilité demandée est donc $\frac{2}{3}$ puisqu'il y a deux tiroirs à pièce d'or dans le premier coffret.

Problème XLI. *On a dix urnes dont l'une* A *renferme* 5 *boules blanches et* 1 *noire, les* 9 *autres* B *renfermant chacune* 2 *blanches et* 2 *noires. On extrait une boule d'une urne choisie au hasard et on constate qu'elle est blanche. Quelle est la probabilité pour que l'urne choisie soit* A ?

On ne peut pas ici raisonner aussi simplement que dans les problèmes précédents parce que les urnes B ne renferment pas le même nombre de boules que l'urne A ; on peut, par l'emploi du plus petit multiple commun, être ramené à ce cas. Rien ne serait changé en effet, si l'on doublait le nombre des boules de l'urne A, en conservant leurs proportions et si l'on triplait de même le nombre de boules des urnes B ; l'urne A renferme alors 10 blanches et 2 noires et chaque urne B renferme 6 blanches et 6 noires ; il y a en tout $10 + 6 \times 9 = 64$ boules blanches parmi lesquelles 10 sont dans A ; la probabilité pour qu'une boule blanche étant extraite, l'urne A ait été choisie est donc $\frac{10}{64} = \frac{5}{32}$.

58. Formule de Bayes. — Considérons n urnes parmi lesquelles n_1 que nous désignerons indifféremment par A_1 ont la même composition et renferment une proportion p_1 de boules blanches ; de même n_2 urnes A_2 renferment une proportion p_2 de boules blanches, ..., n_k urnes A_k renferment une proportion p_k de boules blanches. Les $n = n_1 + n_2 + \ldots + n_k$ urnes étant supposées identiques extérieurement, on choisit l'une d'elles et on en extrait une boule, qui se trouve être blanche. Quelle est la probabilité P_1 pour que l'urne choisie soit l'une des urnes A_1 ?

Nous supposerons que chaque urne renferme le même nombre N de boules ; les urnes A_1 renfermeront chacune $p_1 N$ boules blanches [1] les urnes A_2 en renfermeront chacune $p_2 N$, etc. de sorte que le nombre total de boules blanches est

$$n_1 p_1 N + n_2 p_2 N + \ldots + n_k p_k N$$

parmi lesquelles $n_1 p_1 N$ appartiennent à l'une des urnes A_1. La probabilité P_1 pour que la boule blanche extraite provienne d'une urne A_1 est donc :

$$P_1 = \frac{n_1 p_1 N}{n_1 p_1 N + \ldots + n_k p_k N} = \frac{n_1 p_1}{n_1 p_1 + n_2 p_2 + \ldots + n_k p_k}.$$

On écrit souvent cette formule sous une autre forme en posant :

$$n_1 = \varpi_1 n \quad , \quad n_2 = \varpi_2 n, \quad , \quad \ldots . \; n_k = \varpi_k n.$$

La formule devient alors

$$P_1 = \frac{p_1 \varpi_1}{p_1 \varpi_1 + p_2 \varpi_2 + \ldots p_n \varpi_n}.$$

La signification des nombres ϖ_k est très simple.

Si l'on choisit une urne au hasard, la probabilité pour que l'urne choisie soit de l'espèce A_1 est ϖ_1, puisque ϖ_1 est le rapport du nombre n_1 des urnes A_1 au nombre total n des urnes. On dira pour cette raison, que ϖ_1 est la *probabilité à priori* c'est-à-dire avant tout tirage pour que l'urne choisie soit de l'espèce A_1 : la probabilité que nous avons calculée P_1 est alors dite la probabilité *à posteriori* et la formule précédente est la *formule de Bayes*. On peut la démontrer par le raisonnement suivant, où l'on fait usage des théorèmes sur les probabilités totales et composées.

Evaluons la probabilité pour que l'extraction donne *une boule blanche extraite de* A_1 ; nous évaluerons cette probabilité de deux manières en intervertissant l'ordre des deux conditions

---

[1] Le nombre N est choisi tel que $p_1 N$, $p_2 N$, etc. soient des nombres entiers.

auxquelles doit satisfaire la boule : 1° elle doit être blanche ; 2° elle doit être extraite de A_1.

1° La probabilité pour que la boule extraite soit blanche est une probabilité totale, car la boule blanche peut être extraite, soit des urnes A_1, soit des urnes A_2, etc ; sa valeur est donc :

$$p_1\varpi_1 + p_2\varpi_2 + \ldots + p_n\varpi_n.$$

D'autre part, la probabilité pour qu'une boule extraite, *lorsqu'on sait qu'elle est blanche*, provienne de A_1 a été désignée par P_1 : la probabilité demandée est donc

$$P_1(p_1\varpi_1 + p_2\varpi_2 + \ldots + p_n\varpi_n). \tag{1}$$

2° La probabilité pour que la boule extraite provienne de A_1 est égale à ϖ_1 ; sachant qu'elle vient de A_1, la probabilité pour qu'elle soit blanche est p_1 : la probabilité demandée est donc

$$p_1\varpi_1. \tag{2}$$

En égalant les deux valeurs trouvées (1) et (2), on retrouve bien la formule de Bayes. Cette dernière démonstration s'applique quelle que soit la nature des expériences dans lesquelles les probabilités sont mises en jeu : ce qui en rend parfois les applications difficiles, c'est l'incertitude qui règne sur la valeur des probabilités *à priori*. Donnons en quelques exemples.

Problème XLII. *Une urne A renferme un nombre inconnu de boules blanches ou noires ; on fait plusieurs tirages successifs, en remettant la boule extraite après chaque tirage ; le résultat de ces tirages est l'extraction de r boules blanches et de s boules noires. Quelle est la composition la plus probable de l'urne ?*

On est tenté de répondre que la composition la plus probable est telle que le rapport du nombre des boules blanches au nombre des boules noires est égal au rapport de r à s ; mais ce résultat ne sera exact que si l'on fait sur les probabilités *à priori* des hypothèses non seulement très particulières, mais variables avec les valeurs de r et de s ; c'est ce que montrera l'étude de quelques exemples particuliers.

1° Supposons que l'on ait $r = 2$, $s = 0$ c'est-à-dire que deux tirages consécutifs aient donné des boules blanches, et que l'on sache que l'urne renferme 6 boules. On peut faire sur sa composition 7 hypothèses que nous désignerons par A_0, A_1, ..., A_6, l'hypothèse A_k étant celle où l'urne renferme $6 - k$ blanches et k noires. Dans chacune de ces hypothèses la probabilité d'extraire deux blanches consécutives est donnée par le tableau suivant

A_0	6 blanches 0 noire	$p_0 = 1$
A_1	5 blanches 1 noire	$p_1 = \left(\frac{5}{6}\right)^2 = \frac{25}{36}$
A_2	4 blanches 2 noires	$p_2 = \left(\frac{4}{6}\right)^2 = \frac{16}{36}$
A_3	3 blanches 3 noires	$p_3 = \left(\frac{3}{6}\right)^2 = \frac{9}{36}$
A_4	2 blanches 4 noires	$p_4 = \left(\frac{2}{6}\right)^2 = \frac{4}{36}$
A_5	1 blanche 5 noires	$p_5 = \left(\frac{1}{6}\right)^2 = \frac{1}{36}$
A_6	0 blanche 6 noires	$p_6 = 0$

Si donc on désigne par ϖ_0, ϖ_1, ϖ_2, ... ϖ_6 les probabilités *à priori* de ces 7 hypothèses, leurs probabilités *à posteriori* P_0, P_1, P_2, ..., P_6 sont données par les formules

$$P_0 = \frac{36\varpi_0}{36\varpi_0 + 25\varpi_1 + 16\varpi_2 + 9\varpi_3 + 4\varpi_4 + \varpi_5}$$

$$P_1 = \frac{25\varpi_1}{36\varpi_0 + 25\varpi_1 + 16\varpi_2 + 9\varpi_3 + 4\varpi_4 + \varpi_5}$$

. . . . . .

$$P_5 = \frac{\varpi_5}{36\varpi_0 + 25\varpi_1 + 16\varpi_2 + 9\varpi_3 + 4\varpi_4 + \varpi_5}$$

$$P_6 = 0$$

On voit qu'il est possible de déterminer ϖ_0, ϖ_1, ..., ϖ_5 de manière que P_0, P_1, ..., P_5 aient des valeurs positives *quelconques* dont la somme soit égale à l'unité (on a forcément $P_6 = 0$, car si l'on a extrait des boules blanches, on n'est sûrement pas dans le cas où l'urne ne renferme que des noires).

Il est donc nécessaire de faire des hypothèses sur ϖ_0, ϖ_1, ..., ϖ_7, c'est-à-dire de compléter l'énoncé. Une hypothèse plausible consistera à supposer ces probabilités *à priori* égales entre elles, c'est-à-dire à admettre que les 7 compositions possibles sont représentées par 7 urnes entre lesquelles on choisit au hasard ; cette hypothèse donne :

$$P_0 = \frac{36}{91}, \quad P_1 = \frac{25}{91}, \quad P_2 = \frac{16}{91}, \quad P_3 = \frac{9}{91}, \quad P_4 = \frac{4}{91}, \quad P_5 = \frac{1}{91}.$$

Une autre hypothèse consistera à admettre que l'urne a été remplie au moyen de tirages effectués dans une urne renfermant en proportions égales des boules blanches et noires ; on a alors d'après le triangle arithmétique :

$$\varpi_0 = \varpi_6 = \frac{1}{64}$$

$$\varpi_1 = \varpi_5 = \frac{6}{64}$$

$$\varpi_2 = \varpi_4 = \frac{15}{64}$$

$$\varpi_3 = \frac{20}{64}$$

et par suite, on obtient

$$P_0 = \frac{6}{112}, \quad P_1 = \frac{25}{112}, \quad P_2 = \frac{40}{112}, \quad P_3 = \frac{30}{112}, \quad P_4 = \frac{10}{112}, \quad P_5 = \frac{1}{112}.$$

On voit que ce résultat diffère beaucoup du précédent ; l'hypothèse la plus probable est ici celle où l'urne renferme 4 blanches et 2 noires, parce que la probabilité *à priori* de cette hypothèse est bien plus grande que la probabilité *à priori* de l'hypothèse où l'urne renfermerait exclusivement des boules blanches.

Supposons maintenant que l'on ait $r = s = 1000$, c'est-à-dire que 2000 tirages aient donné égalité entre le nombre des boules blanches et des boules noires extraites ; cherchons la probabilité pour que le rapport des boules blanches au nombre total des boules soit p, le rapport des boules noires au nombre total des

boules étant q. Si ce rapport est réalisé, le nombre normal de blanches extraites sur 2000 tirages sera $2000\,p$ et par suite l'écart absolu observé sera

$$2000\,p - 1000 = 1000\,(2p - 1)$$

l'unité d'écart étant

$$\sqrt{1000\,pq} = 20\sqrt{10\,pq}.$$

L'écart relatif λ est donc

$$\lambda = \frac{1000\,(2p-1)}{10\sqrt{10\,pq}} = 50\,\frac{(2p-1)}{\sqrt{10pq}}$$

et l'on sait que la probabilité d'un tel écart est extrêmement faible lorsque λ dépasse 5 ; on pourra donc la négliger, à moins que la probabilité *à priori* correspondante ne soit très élevée. Prenons donc

$$p = \frac{1}{2} + \alpha$$

$$q = \frac{1}{2} - \alpha$$

et négligeons le carré de α ; il vient

$$\lambda = \frac{200\,\alpha}{\sqrt{10}} = 20\,\alpha\sqrt{10}$$

et λ dépasse $\sqrt{10}$ lorsque α dépasse $\frac{1}{20}$. Supposons que la probabilité *à priori* pour que α ait une valeur comprise dans un intervalle déterminé compris entre $-\frac{1}{20} + \frac{1}{20}$ soit proportionnelle à l'étendue de cet intervalle : la probabilité pour que, α étant compris entre α et $\alpha + d\alpha$, on observe un écart relatif λ compris entre $20\alpha\sqrt{10}$ et $20\alpha\sqrt{10} + d\lambda$ est

$$\frac{1}{\sqrt{\pi}}\,e^{-1000\alpha^2}\,d\lambda.$$

La probabilité pour que z soit compris entre z et $z + dz$ est donc

$$\frac{\frac{1}{\sqrt{\pi}} e^{-1000z^2} dz}{\frac{1}{\sqrt{\pi}} \int_{-\frac{1}{20}}^{+\frac{1}{20}} e^{-1000z^2} dz} = \sqrt{1000}\, e^{-1000z^2} dz$$

car l'intégrale du dénominateur peut être évaluée sans erreur sensible entre $-\infty$ et $+\infty$.

La probabilité pour que le rapport entre le nombre des boules blanches et noires soit compris entre 0,5 et 0,501 s'obtiendra en prenant $z = 0$, $dz = \frac{1}{1000}$ ce qui donne

$$\frac{\sqrt{1000}}{1000} = \frac{2\sqrt{10}}{100}$$

la probabilité pour que ce rapport soit compris entre 0,510 et 0,511 s'obtiendra en prenant $z = \frac{1}{100}$, $dz = \frac{1}{1000}$ ce qui donne

$$\frac{\sqrt{1000}}{1000} e^{-0,1}.$$

Enfin, la probabilité pour que ce rapport soit compris entre 0,525 et 0,526 s'obtiendra en prenant $z = \frac{1}{40}$, $dz = 0,001$, ce qui donne

$$\frac{\sqrt{4000}}{1000} e^{-2,5}$$

elle est environ 10 fois plus faible que la précédente.

3° Supposons que l'on ait $r = 10$, $s = 0$ et que l'urne ait été remplie au moyen de 800 tirages effectuées dans une autre urne, où les nombres de boules blanches et noires étaient égaux. La probabilité *à priori* pour que l'urne renferme $500 + h$ boules blanches est dans ce cas

$$\frac{1}{40\sqrt{\pi}} e^{-\left(\frac{h}{20}\right)^2}$$

car l'unité d'écart est égale à 20. Nous nous bornerons à rechercher la composition *la plus probable* : la probabilité pour que 10 tirages consécutifs donnent des boules blanches est

$$\left(\frac{400+h}{800}\right)^{10},$$

de sorte que nous sommes amenés à rechercher le maximum du produit

$$\frac{1}{40\sqrt{\pi}}e^{-\left(\frac{h}{40}\right)^2}\left(\frac{400+h}{800}\right)^{10}$$

dont le logarithme est, à une constante près

$$10\log\left(1+\frac{h}{400}\right)-\frac{h^2}{400},$$

Prenons la dérivée, nous obtenons

$$\frac{10}{400+h}-\frac{h}{200}=0$$

$$h^2+400h-2000=0$$

$$h=-200+\sqrt{42000}=200\left[-1+\sqrt{1+\frac{1}{20}}\right]$$

$$=5 \text{ environ.}$$

Ainsi, la composition la plus probable est 405 boules blanches et 395 boules noires ; l'extraction de 10 boules blanches sur 10 tirages ne rend pas très vraisemblable un écart plus considérable dans la composition de l'urne ; ceci tient à ce que, d'après la manière dont l'urne a été composée, il est extrêmement peu probable que le nombre de boules blanches soit grand par rapport au nombre des noires.

Il en serait autrement si 100 tirages consécutifs donnaient 100 boules blanches ; dans ce cas, on devrait tout d'abord se demander s'il n'y a pas eu quelque erreur d'expérience car un tel résultat serait extrêmement invraisemblable ; si on l'admet et

que l'on recommence le calcul on est amené à chercher le maximum de l'expression

$$100 \log (400 + h) - \frac{h^2}{400}$$

ce qui donne

$$\frac{100}{400 + h} - \frac{h}{200} = 0$$

$$h^2 + 400\,h - 20\,000 = 0$$

$$h = -200 + \sqrt{60\,000} = 44 \text{ environ.}$$

59. Applications diverses. — Il est des cas où l'évaluation de la probabilité *à priori* est une simple affaire d'appréciation : le résultat est fonction de la valeur qui est attribuée à cette probabilité. En voici un exemple classique.

Problème XLIII. *Pierre joue à l'écarté avec un inconnu qui, la première fois qu'il donne les cartes, retourne le roi. Quelle est la probabilité pour que le partenaire de Pierre soit tricheur de profession ?*

Dans la solution que M. Henri Poincaré donne de ce problème, il admet implicitement qu'un tricheur de profession tourne le roi chaque fois qu'il en a l'occasion. Bien que je n'aie pas — plus que M. Poincaré — fréquenté « le monde où l'on triche » cette hypothèse me paraît assez peu vraisemblable ; un joueur honnête tourne le roi en moyenne une fois sur huit : il paraît difficile qu'un individu à qui pareille bonne fortune arriverait à chaque coup n'attire pas rapidement l'attention ; un tricheur, d'ailleurs, préférera varier ses moyens d'action, et, en tous cas, ne cherchera pas à forcer le sort lorque celui-ci lui sera spontanément favorable. Il semble donc que l'on se rapproche de la vérité en évaluant à $\frac{1}{4}$, par exemple, la probabilité pour qu'un tricheur retourne le roi, alors que cette probabilité est $\frac{1}{8}$ pour le joueur honnête.

Désignons par ϖ la probabilité *à priori* pour que le partenaire de Pierre soit tricheur : la probabilité *à priori* pour qu'il soit honnête est $1 - \varpi$ et la probabilité P *à posteriori* c'est-à-dire après qu'il a tourné le roi pour qu'il soit tricheur est

$$P = \frac{\frac{1}{4}\varpi}{\frac{1}{4}\varpi + \frac{1}{8}(1-\varpi)} = \frac{2\varpi}{1+\varpi}.$$

Si $\varpi = \frac{1}{n}$, c'est-à-dire si l'on suppose égales les probabilités *à priori* pour que l'adversaire de Pierre soit tricheur et pour qu'il ne le soit pas, on trouve

$$P = \frac{2}{3}$$

c'est-à-dire que l'unique événement observé augmente notablement la probabilité. Mais l'on observe justement que Pierre serait bien imprudent d'engager une partie contre un partenaire de la moralité spéciale duquel il aurait une aussi fâcheuse opinion : le fait que Pierre joue suppose qu'il regarde ϖ comme un nombre assez petit. Si ϖ est petit, P diffère peu de 2ϖ ; en tous cas d'ailleurs, P est inférieur à 2ϖ, c'est-à-dire que la probabilité est tout au plus doublée par l'événement observé ; si le partenaire de Pierre est tel que ϖ soit égal à 1 sur 100 000, P est égal à 1 sur 50 000, c'est-à-dire pratiquement aussi négligeable que ϖ ; si Pierre est sûr de son partenaire et suppose $\varpi = 0$, on a aussi $P = 0$, c'est-à-dire que l'événement ne doit pas ébranler sa confiance : les résultats du calcul sont conformes à ceux du bon sens.

Problème XLIV. *Un examinateur ayant préparé un choix de dix questions écrites en fait tirer au sort une par chaque candidat. De deux candidats, le plus instruit connaît 9 questions sur 10 et l'autre seulement 1 sur 10 ; un de ces candidats tire une question qu'il connaît : quelle est la probabilité pour que ce soit le plus instruit ?*

La probabilité *à priori* est $\frac{1}{2}$, la probabilité *à posteriori* est donc :

$$P = \frac{\frac{1}{2}\cdot\frac{9}{10}}{\frac{1}{2}\cdot\frac{9}{10} + \frac{1}{2}\cdot\frac{1}{10}} = \frac{9}{10}.$$

Problème XLV. *Deux classes de 20 élèves renferment, la première 10 bons élèves, 5 médiocres et 5 mauvais et la seconde 5 bons élèves, 5 médiocres et 10 mauvais. Un inspecteur interroge au hasard un élève de chaque classe et trouve l'élève de la classe* A *supérieur à l'élève de la classe* B *; quelle est la probabilité pour que la classe* A *soit la première ?*

En interrogeant au hasard un élève de la première classe et un élève de la seconde, il y a

$$10 \times 15 + 5 \times 10 = 200$$

combinaisons dans lesquelles l'élève de la première est meilleur ; il y en a

$$10 \times 5 + 5 \times 5 + 5 \times 10 = 125$$

dans lesquelles les deux élèves font partie de la même catégorie ; et enfin, il y a

$$5 \times 10 + 5 \times 5 = 75$$

combinaisons dans lesquelles l'élève de la seconde se trouve meilleur. La probabilité *à priori* étant $\frac{1}{2}$ pour chaque, la probabilité pour que la classe A soit la première est

$$\frac{\frac{1}{2} \times 200}{\frac{1}{2} \times 200 + \frac{1}{2} \times 75} = \frac{8}{11}.$$

Dans le cas où l'on aurait interrogé deux élèves de chaque classe et où chaque élève de la classe A se serait montré supé-

rieur à l'élève correspondant de B, la probabilité pour que la classe A soit la première serait

$$\frac{\frac{1}{2}(200)^2}{\frac{1}{2}(200)^2 + \frac{1}{2}(75)^2} = \frac{8^2}{8^2 + 3^2} = \frac{64}{73}.$$

On voit combien la répétition des épreuves augmente les chances d'aboutir à une conclusion exacte, en supposant, bien entendu, qu'on ne commette pas d'erreurs dans le jugement porté sur les individus comme nous l'avons implicitement supposé dans l'énoncé de ce problème.

Problème XLVI. *Un régiment renferme des soldats de première, de seconde et de troisième année ; on y choisit deux soldats au hasard et l'un d'eux déclare qu'il est plus ancien que l'autre ; quelle est la probabilité pour qu'il soit de troisième année ?*

Soit a le nombre des soldats de première année, b le nombre des soldats de seconde, c le nombre des soldats de troisième. Etant donné un soldat de seconde année et un autre pris au hasard, la probabilité pour que le premier soit plus ancien est

$$\frac{a}{a + b + c - 1}$$

car le second peut être l'un quelconque des $a + b + c - 1$ autres que le premier et sur ces hypothèses, a seulement sont favorables. De même, la probabilité pour qu'un soldat de troisième année soit plus ancien qu'un autre pris au hasard est

$$\frac{a + b}{a + b + c - 1}.$$

D'autre part, les probabilités *à priori* pour qu'un soldat soit de seconde ou de troisième année sont respectivement [1]

$$\frac{b}{a + b + c}, \qquad \frac{c}{a + b + c}.$$

[1] Il n'y a pas lieu de s'occuper des cas où le soldat est de première année car il ne peut alors être plus ancien que le second.

La probabilité *à posteriori* pour que le soldat le plus ancien soit de troisième année est donc

$$\frac{\frac{c}{a+b+c}\times\frac{a+b}{a+b+c-1}}{\frac{c}{a+b+c}\times\frac{a+b}{a+b+c-1}+\frac{b}{a+b+c}\times\frac{a}{a+b+c-1}}$$

c'est-à-dire

$$\frac{c(a+b)}{c(a+b)+ab}=1-\frac{ab}{bc+ca+ab}.$$

On peut aussi écrire cette probabilité sous la forme

$$\frac{\frac{1}{a}+\frac{1}{b}}{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}}.$$

Si $a=b=c$ la probabilité est $\frac{2}{3}$; si $a=2b=4c$, c'est-à-dire si chaque année est moitié moins nombreuse que la précédente, la probabilité devient

$$\frac{\frac{1}{4}+\frac{1}{2}}{\frac{1}{4}+\frac{1}{2}+1}=\frac{3}{7}$$

c'est-à-dire est inférieure à $\frac{1}{2}$; dans ce cas, il y aurait avantage à parier que le soldat le plus ancien est de seconde année, parce que les soldats de seconde année sont deux fois plus nombreux que ceux de troisième et ceux de première année deux fois plus nombreux que ceux de seconde. Cet exemple montre nettement l'influence de la probabilité *à priori*.

CHAPITRE XI

PROBLÈMES STATISTIQUES
SCHÉMA DES URNES

60. Notion de la probabilité statistique. — L'étude de certaines statistiques montre qu'il existe entre les évaluations numériques de divers phénomènes des rapports sensiblement constants; par exemple, le rapport du nombre des naissances au nombre des habitants ne varie pas beaucoup à une époque donnée, dans une région donnée, lorsqu'on l'observe pendant plusieurs années consécutives. Un tel phénomène n'est nullement comparable *à priori* à un tirage effectué dans une urne renfermant une proportion déterminé de boules blanches et noires ; la seule analogie qui apparaisse immédiatement entre ces deux phénomènes, c'est la constance d'un certain rapport. On a été amené à évaluer ce rapport, et bien d'autres du même genre, d'une manière suffisamment approchée, et on leur a donné le nom de *probabilités statistiques*. Il ne faut pas perdre de vue qu'il y a la même différence entre ces probabilités statistiques et les probabilités abstraitement et rigoureusement définies qu'entre les figures étudiées en géométrie et les représentations plus ou moins grossières qu'on en rencontre dans la nature ; entre une sphère, par exemple, et une orange.

61. Naissances masculines et féminines. — L'une des probabilités statistiques les plus anciennement étudiées, et dont la constance est la plus remarquable, est *la probabilité pour qu'un nouveau-né appartienne au sexe masculin*. On sait depuis fort

longtemps que cette probabilité est légèrement supérieure à $\frac{1}{2}$, c'est-à-dire que le nombre des naissances des garçons est légèrement supérieur au nombre des naissances des filles. Nous ne résumerons pas ici toutes les recherches qui ont été faites sur cette question, nous nous contenterons d'indiquer comment on doit poser rationnellement le problème et en rechercher la solution.

Supposons que l'on ait constaté que, pendant une longue période, dans un pays déterminé, le rapport du nombre des naissances des garçons au nombre total des naissances [1] a été $\frac{1}{2} + \alpha$, le rapport du nombre des naissances des filles au nombre total des naissances étant par suite $\frac{1}{2} - \alpha$; le rapport du nombre des naissances masculines au nombre des naissances féminines est par uite

$$\frac{\frac{1}{2} + \alpha}{\frac{1}{2} - \alpha} = 1 + 4\alpha + 8\alpha^2 + \dots$$

La valeur de α est généralement voisine de 0,01 et ne dépasse pas 0,015; le terme en α^2 peut généralement être négligé au degré d'approximation des observations dont on dispose.

La question qui se pose tout d'abord est la suivante : le coefficient α ayant été calculé au moyen d'une série d'observations portant, par exemple, sur 20 ans, on peut, pour chaque année de cette période, calculer *ce qu'aurait dû être le nombre des naissances masculines* en multipliant par $\frac{1}{2} + \alpha$ le nombre total des naissances. Le nombre ainsi calculé différera généralement du nombre observé; *l'écart satisfait-il aux mêmes lois que si chaque naissance était remplacée par un tirage dans une urne renfermant* N *boules parmi lesquelles* $\left(\frac{1}{2} + \alpha\right)$ N *seraient*

[1] Les chiffres varient suivant que l'on fait entrer ou non en compte les morts-nés; ils ne sont pas absolument les mêmes pour les naissances légitimes et pour les naissances illégitimes. Il semble prudent, surtout pour les statistiques anciennes, de se borner aux naissances *légitimes* d'enfants *vivants*, car elles paraissent présenter plus de garanties d'exactitude.

blanches ? Si dans une telle urne on fait n tirages, le nombre le plus probable de boules blanches extraites est

$$\left(\frac{1}{2} + \alpha\right) n$$

et l'unité d'écart a pour valeur

$$\sqrt{2npq} = \sqrt{2n\left(\frac{1}{2} + \alpha\right)\left(\frac{1}{2} - \alpha\right)} = \sqrt{\frac{n}{2} - 2n\alpha^2}.$$

Le carré de α étant très petit, on pourra, sans erreur sensible, négliger le terme en α^2.

On calculera l'écart relatif λ en divisant l'écart absolu par l'unité d'écart et on recherchera si les écarts relatifs se répartissent suivant la loi de Gauss.

Nous nous contenterons d'indiquer ici que la vérification est assez satisfaisante, d'où l'on conclut que, *au moins à une première approximation*, l'assimilation à un tirage dans une urne est légitime [1].

Nous admettrons provisoirement ce résultat, et nous allons voir comment d'autres observations statistiques conduisent à le modifier.

62. Etude des accouchements doubles. — On a observé depuis longtemps que deux jumeaux sont plus fréquemment du même sexe que deux enfants non jumeaux d'une même famille. Peut-on soumettre la question au calcul? L'observation statistique fait connaître les nombres d'accouchements doubles ayant produit soit deux garçons (GG), soit deux filles (FF), soit un garçon et une fille (GF); on en conclut les probabilités statistiques de ces trois hypothèses :

$$\begin{gathered} \text{GG} : a \\ \text{FF} : b \\ \text{GF} : c \\ a + b + c = 1. \end{gathered}$$

---

(1) J'ai discuté cette question et signalé quelques-unes des restrictions qui s'imposent dans mon cours de 1907-1908.

Les nombres a, b, c étant connus, on peut résoudre, en se plaçant uniquement au point de vue des faits déjà observés (c'est-à-dire du passé et non de l'avenir), des problèmes tels que le suivant.

PROBLÈME XLVII. *Sachant que l'un des deux jumeaux est un garçon, quelle est la probabilité pour que l'autre soit aussi un garçon?*

Ce problème est tout à fait semblable au problème des trois coffrets (Problème XL, n° **57**), sauf que nous avons ici aN cas GG et cN cas GF au lieu d'un seul coffret or-or et d'un seul coffret or-argent : la probabilité est donc

$$\frac{2a}{2a + c}.$$

Mais ce n'est pas là le problème dont la solution est la plus intéressante : la question qui se pose naturellement est de savoir dans quel cas le sexe de l'un des deux enfants peut être regardé comme ayant pour *cause* le sexe de l'autre. La principale difficulté est de formuler cette question sous une forme précise et claire : il est nécessaire pour cela, de lui donner un énoncé entièrement objectif, c'est-à-dire indépendant de toute interprétation physiologique ; on peut le faire de bien des manières, entre lesquelles le choix est arbitraire ; nous adopterons la suivante.

PROBLÈME XLVIII. *On fait un certain nombre de séries de deux tirages dans une urne renfermant dans une proportion déterminée, des boules marquées* G *et des boules marquées* F ; *chaque tirage est noté par un aide qui inscrit, suivant les cas,* GG, GF, FG, FF : *mais on suppose que l'aide est parfois inattentif, et que, dans ce cas, au lieu d'écouter le nom de la seconde boule extraite, il inscrit la même lettre que pour la première. Connaissant les résultats globaux des inscriptions faites par l'aide, on demande de déterminer la proportion des cas où il a été inattentif?*

Nous admettrons que les résultats sont donnés par le tableau :

GG	aN
FF	bN
GF + FG	cN

en désignant par N le nombre total des tirages. En désignant par $\frac{1}{2} + \alpha$ la proportion des boules G et par λ la proportion des cas où l'aide a été inattentif, on a les trois relations

$$\begin{cases} (1-\lambda)\left(\frac{1}{2}+\alpha\right)^2 + \lambda\left(\frac{1}{2}+\alpha\right) = a \\ (1-\lambda)\left(\frac{1}{2}-\alpha\right)^2 + \lambda\left(\frac{1}{2}-\alpha\right) = b \\ 2(1-\lambda)\left(\frac{1}{2}+\alpha\right)\left(\frac{1}{2}-\alpha\right) = c \end{cases}$$

qui se réduisent à deux, car on a identiquement

$$a + b + c = 1.$$

Si l'on suppose α donné et λ seul inconnu, le problème ne sera possible que si les nombres a, b, c, satisfont à une condition, qui s'obtient aisément en retranchant les deux premières équations membre à membre, ce qui donne :

$$2\alpha = a - b.$$

On constate que cette valeur de α coïncide d'une manière suffisamment exacte avec la valeur de α déduite de l'ensemble des naissances ; le système peut donc être réduit à la première équation, ou à la seconde ; on aura une équation plus symétrique en les ajoutant membre à membre, ce qui donne

$$(1-\lambda)\left(\frac{1}{2} - 2\alpha^2\right) + \lambda = a + b$$

ou, en négligeant α^2 :

$$\lambda = 2a + 2b - 1.$$

En tenant compte de la relation

$$a + b + c = 1$$

on peut écrire

$$\lambda = a + b - c = 1 - 2c.$$

En pratique, la formule la plus commode sera la suivante :

$$\lambda = \frac{aN + bN - cN}{N}.$$

Par exemple en 1905, la statistique de la France donne les chiffres suivants :

GG	$aN = 2933$
FF	$bN = 2797$
GF	$cN = 3188.$

On en conclut :

$$\lambda = \frac{2542}{8918} = 0{,}28\ldots$$

On peut donc admettre que tout se passe comme si 28 fois sur cent environ, le sexe du second enfant était déterminé par le sexe du premier.

63. Naissances dans une même famille. — Une recherche analogue pourrait être faite sur les naissances consécutives dans une même famille ; malheureusement les documents statistiques sont trop peu nombreux pour permettre une étude approfondie de la question. On peut utiliser de la manière la plus avantageuse les quelques statistiques que l'on possède en faisant usage de l'artifice suivant [1]. En supposant que la probabilité de la nais-

(1) Voir dans les documents relatifs à la session de 1908 du Conseil supérieur de statistique, le rapport sur la statistique des fonctionnaires faite sous la direction de M. Lucien March, où la méthode que j'indique a été appliquée pour la première fois. Je tiens à remercier M. Lucien March de l'obligeance avec laquelle il a mis à ma disposition pour cette recherche, faite une statistique sous sa direction à la statistique générale de la France.

sance d'un garçon soit toujours $\frac{1}{2} + \alpha$, les probabilités des quatre hypothèses GG, GF, FG, FF seront respectivement

$$\text{GG} \quad \left(\frac{1}{2} + \alpha\right)^2 = \frac{1}{4} + \alpha + \alpha^2$$

$$\text{GF} \quad \left(\frac{1}{2} + \alpha\right)\left(\frac{1}{2} - \alpha\right) = \frac{1}{4} - \alpha^2$$

$$\text{FG} \quad \left(\frac{1}{2} - \alpha\right)\left(\frac{1}{2} + \alpha\right) = \frac{1}{4} - \alpha^2$$

$$\text{FF} \quad \left(\frac{1}{2} - \alpha\right)^2 = \frac{1}{4} - \alpha + \alpha^2.$$

On en conclut :

$$\frac{\text{GG} + \text{FF}}{\text{GF} + \text{FG}} = \frac{\frac{1}{2} + 2\alpha^2}{\frac{1}{2} - 2\alpha^2} = \frac{1 + 4\alpha^2}{1 - 4\alpha^2} = 1 + 8\alpha^2 + \dots$$

La valeur de α étant voisine de 0,01, le rapport ainsi calculé doit donc différer de l'unité de moins de 0,001 ; de plus, si l'on suppose que la population sur laquelle porte la statistique, se compose de plusieurs catégories pour lesquelles α a diverses valeurs $\alpha_1, \alpha_2, \dots \alpha_k$, la valeur du rapport calculé sera, en désignant par $n_1, n_2, \dots, n_k$ les nombres respectifs d'observations dans chaque catégorie

$$\frac{n_1(1 + 4\alpha^2_1) + n_2(1 + 4\alpha^2_2) + \dots + n_k(1 + 4\alpha^2_k)}{n_1(1 - 4\alpha^2_1) + n_2(1 - 4\alpha^2_2) + \dots + n_k(1 - 4\alpha^2_k)}$$

et par suite différera de l'unité d'une quantité moindre que $8\alpha^2$, α désignant le plus grand des nombres $\alpha_1, \alpha_2, \dots \alpha_k$,

Si donc le rapport calculé :

$$\frac{\text{GG} + \text{FF}}{\text{GF} + \text{FG}}$$

est sensiblement supérieur à l'unité, on pourra conclure avec quelque *vraisemblance* qu'il y a des cas où le sexe du second

enfant est déterminé par le sexe du premier. Nous n'insisterons pas sur l'évaluation de cette vraisemblance, car, nous le répétons, les documents statistiques ne sont pas assez étendus pour que l'on puisse opérer sur des nombres suffisamment grands pour écarter avec certitude les petits écarts fortuits, toujours possibles sur de petits nombres.

64. Statistique des décès. — Lorsque l'on examine, même superficiellement, la statistique des décès survenus dans une région donnée [1], pendant plusieurs années consécutives, on est immédiatement frappé par la grandeur relativement considérable des écarts. Par exemple, en France, les chiffres des décès ont été les suivants dans les 20 dernières années.

année	mille	année	mille
1887	843	1897	751
1888	838	1898	810
1889	795	1899	816
1890	877	1900	853
1891	877	1901	785
1892	876	1902	761
1893	868	1903	754
1894	816	1904	761
1895	852	1905	770
1896	772	1906	780

Comment se rendre compte de ces écarts considérables ? Une première explication qui vient naturellement à l'esprit se rattache au fait que la population sur laquelle porte la statistique n'est pas homogène : la probabilité du décès dépend évidemment de l'âge, du sexe, de la situation sociale, du lieu de résidence, etc. Mais nous savons (n^os^ **32. 33. 34**) que la considération simultanée de plusieurs groupes distincts, que l'on regarde comme n'en

(1) Si la population de la région étudiée n'est pas sensiblement constante pendant la période de temps considérée, on aura soin de réduire par un calcul proportionnel les chiffres des décès à une population constante égale à la moyenne.

formant qu'un, ne peut que *diminuer* l'écart relatif et ne l'augmente jamais. Il faut donc chercher une autre explication : elle nous sera suggérée par l'observation des faits : le nombre des décès est fréquemment accru par une épidémie, par une circonstance climatérique comme l'excès du froid ou de la chaleur. De tels phénomènes ont pour résultat de faire mourir simultanément un assez grand nombre d'individus ; ces décès ne doivent donc pas être regardés comme indépendants les uns des autres : au point de vue purement abstrait, tout se passe comme s'ils étaient représentés par des boules liées entre elles, de telle sorte que l'on ne puisse pas extraire l'une d'elles d une urne sans extraire toutes les autres. Or, ceci justifie l'existence de plus grands écarts. Supposons, pour fixer les idées, qu'une urne renferme 50 000 000 de boules, parmi lesquelles 800 000 sont noires ; 50 millions de personnes sont invitées à extraire chacune une boule, aussitôt remise dans l'urne ; le nombre total probable des boules noires extraites sera évidemment 800 000 ; l'unité d'écart sera pour $x = 50\,000\,000$ $p = \frac{1}{50}$ $q = \frac{49}{50}$

$$\sqrt{2npq} = \frac{1\,000}{50} \cdot 7\sqrt{80} = 1\,260 \text{ environ.}$$

Supposons maintenant que l'on fasse seulement 50 000 extractions, mais que chaque extraction d'une boule noire soit considérée comme en valant 1 000 ; le nombre total probable de boules noires extraites sera 800 nombre réel et par suite 800 000 d'après la convention faite ; l'unité d'écart réelle sera

$$\sqrt{800\,000 \cdot \frac{1}{50} \cdot \frac{49}{50}} = 7 \cdot \frac{100}{50}\sqrt{8} = 28\sqrt{2} = 40 \text{ environ.}$$

mais, d'après la convention faite, chaque unité vaut 1 000 : l'unité d'écart sera donc, en tenant compte de cette convention, équivalente à 50 000, c'est-à-dire qu'un écart de 50 000 sur 800 000 devra être considéré comme assez vraisemblable. On voit que l'hypothèse faite permet de rendre compte d'écarts considérables ; il suffira, dans chaque cas particulier, de la préciser convena-

blement. Laissant de côté la question de la statistique des décès, par laquelle nous avons été conduits aux réflexions précédentes, nous allons examiner brièvement comment on peut poser le problème général de la probabilité statistique.

65. Schéma des urnes. — Pour ne pas introduire d'inutiles complications verbales, nous admettrons que l'on considère des séries d'observations portant chacune sur le même nombre de cas possibles (par exemple sur une population constante) ; dans chaque série, on note le nombre des cas où s'est produit l'événement sur lequel porte l'observation : c'est ce qu'on appellera les *cas favorables*, cette appellation étant consacrée par l'usage (ce peuvent être les nombres des naissances, ou des décès, ou de personnes atteintes de telle maladie, etc.). L'observation fournit donc un tableau indiquant pour chaque série, le nombre des cas favorables ; on connaît d'autre part, le nombre des cas possibles. Le problème général de la statistique mathématique est le suivant :

Déterminer un système de tirages effectués dans des urnes de composition fixe, de telle manière que les résultats d'une série de tirages, interprétés à l'aide de coefficients fixes convenablement choisis, puissent avec une très grande vraisemblance conduire à un tableau identique au tableau des observations.

Nous donnerons au problème ainsi posé le nom de *problème du schéma des urnes*. Le type de la solution qu'il peut recevoir est le suivant : on considère trois urnes U_1, U_2, U_3, renfermant des boules identiques de forme, noires et blanches, la proportion des boules noires étant respectivement p_1, p_2, p_3 ; on fait dans l'urne U_1 un nombre d'extractions égale à α_1 en remettant chaque fois la boule extraite, qui donnent n_1 boules noires ; on extrait de même, en α_2 tirages, n_2 boules noires de U_2 et, en α_3 tirages, n_3 boules noires de U_3 ; le nombre des cas favorables est alors représenté par la somme

$$\lambda_1 n_1 + \lambda_2 n_2 + \lambda_3 n_3$$

les coefficients λ_1, λ_2, λ_3 ainsi que les nombres entiers α_1, α_2, α_3, et les probabilités p_1, p_2, p_3 définissent le schéma d'urnes considéré.

Nous devons nous contenter d'indiquer ici les diverses questions qui se posent à propos du problème général du schéma des urnes.

1° *Ce problème est-il possible ?* Comme il s'agit seulement d'un problème d'approximation, il est clair qu'il est toujours possible en supposant le nombre des urnes suffisamment grand ; mais une telle solution ne serait pas très satisfaisante ; pratiquement, on doit chercher à le résoudre avec des urnes en nombre minimum ; dans les cas où l'on peut y parvenir avec une seule urne, le problème considéré est dit *normal* ; un cas intéressant est celui où deux urnes sont nécessaires et deux seulement.

2° *Le problème admet-il plusieurs solutions ?* Certainement, si l'on suppose le nombre des urnes assez grand ; mais, dans bien des cas, il n'y a qu'une seule *solution simple*, et c'est à celle-là qu'on s'attache particulièrement.

3° *Par quelle méthode obtenir les solutions ?* Je me contenterai de renvoyer à ce sujet aux travaux de Karl Pearson, bien qu'il ne pose pas le problème sous la forme que nous lui avons donnée : la question me paraît d'ailleurs appeler de nouvelles recherches.

Pour terminer ces rapides indications, je tiens à signaler l'analogie entre la forme que nous avons donnée au problème du schéma des urnes et la forme sous laquelle se posent les problèmes de mécanique et de physique mathématique : là, aussi, la première question est de construire un schéma mathématique présentant avec la réalité d'assez étroits rapports. Cette première question résolue, ou du moins partiellement résolue (car il y a souvent plusieurs solutions), le rôle du mathématicien se borne à étudier les propriétés du schéma obtenu, ce qui est un problème de mathématiques pures ; la comparaison des résultats obtenus ainsi avec l'expérience et le développement des théories que peut suggérer cette comparaison sont en dehors du domaine

des mathématiques ; car, on doit, dans ces recherches théoriques, ne jamais perdre de vue les réalités et contrôler à chaque instant les idées nouvelles par l'observation et l'expérience. Mais le rôle des mathématiques, pour être limité, n'en est pas moins, dans bien des cas, fort important.

CHAPITRE XII

CAS DES PROBABILITÉS CONTINUES

66. Etude d'un problème simple. — On joue à un jeu analogue à *pile ou face*, en se servant d'un solide polyédrique dont une ou plusieurs faces ont été peintes en noir, les autres restant blanches : ce solide est jeté en l'air et l'on examine s'il repose finalement sur le sol plan par une face noire ou par une face blanche : ces deux cas correspondant respectivement aux cas de *pile ou face*. La probabilité de *pile* est donc inconnue, puisque l'on ne sait rien du nombre ni de la dimension des faces noires du polyèdre : on ne sait rien non plus du polyèdre lui-même ; ce pourrait être un dé cubique dont cinq faces sur six seraient noires : ce pourrait être aussi un cube de bois dans un angle duquel on aurait taillé une très petite face triangulaire peinte en noir, tout le reste de la surface étant blanc.

Le problème que nous allons étudier est le suivant :

PROBLÈME XLIX. *On sait que $n - p$ parties ont donné p fois pile et n fois face : que peut-on en conclure au sujet de la probabilité x d'amener pile au jeu considéré ?*

D'après les hypothèses, on ne sait rien *à priori* sur cette probabilité x ; *nous admettrons que, dans cette ignorance, on a le droit de considérer comme également vraisemblables toutes les valeurs de x comprises entre* 0 *et* 1. Dès lors, la probabilité *a priori* pour que x soit compris entre x_1 et x_2 est égal à $x_2 - x_1$;

la probabilité pour que x soit compris entre z et $z + dz$ est égale à dz. Lorsque x a la valeur z, la probabilité d'amener p fois pile et n fois face est égale à

$$\frac{n!\,p!}{(n+p)!}\,z^p(1-z)^n \tag{1}$$

car la probabilité d'amener pile étant z, la probabilité d'amener face est $1 - z$. Si nous conservons les notations du n° **58**, les nombres ϖ_k (probabilités *à priori*) sont égaux à dz, et les nombres p_k sont égaux à l'expression (1). Dans l'application de la formule de Bayes

$$P = \frac{p_k \varpi_k}{\Sigma p_k \varpi_k}$$

on peut supprimer le facteur indépendant de z qui figure dans tous les p_k; on obtient ainsi

$$P = \frac{z^p(1-z)^n\,dz}{\int_0^1 z^p(1-z)^n\,dz}.$$

Telle est la probabilité *à posteriori*, c'est-à-dire en tenant compte des événements observés, pour que x soit compris entre z et $z + dz$. Quelles que soient les hypothèses, cette probabilité est infiniment petite en même temps que dz, c'est-à-dire que chaque hypothèse particulière précise sur x est infiniment peu vraisemblable : on doit considérer un intervalle fini pour que la probabilité pour que x soit compris dans cet intervalle ait elle-même une valeur finie. Nous allons appliquer la formule obtenue à quelques cas particuliers.

67. Cas particuliers.

Problème L. *Au jeu de pile ou face défini au paragraphe précédent, deux expériences ont donné toutes deux pile. Quelle somme peut-on parier contre* 1 *franc, que ce résultat est dû au fait que le polyèdre utilisé dans le jeu donne à pile une probabilité plus grande qu'à face ?*

Nous devons supposer $p = 2$, $n = 0$, ce qui donne pour la probabilité élémentaire la valeur

$$P = \frac{z^2 dz}{\int_0^1 z^2 dz} = 3z^2 dz.$$

Pour avoir la probabilité pour que x soit supérieur à $\frac{1}{2}$, nous devons intégrer entre les limites $\frac{1}{2}$ et 1, ce qui donne

$$\int_{\frac{1}{2}}^{1} 3z^2 dz = 1 - \frac{1}{8} = \frac{7}{8}.$$

La probabilité pour que le polyèdre employé donne à pile une probabilité supérieure à $\frac{1}{2}$ est donc $\frac{7}{8}$, la probabilité contraire est par suite $\frac{1}{8}$ et l'on peut parier 7 francs contre 1 franc que l'on se trouve dans le premier cas.

Nous reviendrons sur ce résultat après l'étude du problème LI.

Problème LI. *Les conditions étant les mêmes que dans le problème* XLIX, *on suppose les nombres n et p très grands et l'on pose*

$$\frac{n}{n+p} = t.$$

Quelle est la probabilité pour que x soit compris entre $t - y$ et $t + y$, y, étant un nombre très petit ?

On observera que la valeur t définie dans l'énoncé est la valeur qu'il faut donner à z pour rendre maximum le produit

$$z^n (1 - z)^p$$

et par suite la probabilité élémentaire P. On obtient en effet, en égalant à zéro la dérivée logarithmique du produit précédent

$$\frac{n}{z} - \frac{p}{1 - z} = 0$$

c'est-à-dire

$$\frac{z}{n} = \frac{1 - z}{p} = \frac{1}{n + p}.$$

Evaluons d'autre part l'intégrale

$$J_{n,p} = \int_0^1 z^n (1 - z)^p dz.$$

La formule

$$\frac{d}{dz}\left[z^{n+1}(1 - z)^p\right] = (n + 1) z^n (1 - z)^p - p z^{n+1} (1 - z)^{p-1}$$

donne, par intégration entre les limites 0 et 1 :

$$0 = (n + 1) J_{n,p} - p J_{n+1, p-1}$$

car l'intégrale du premier membre est nulle du moment que $n + 1$ et p sont positifs. La formule précédente s'écrit

$$J_{n,p} = \frac{p}{n + 1} J_{n+1,p-1}$$

et par suite,

$$J_{n,p} = \frac{p}{n + 1} \cdot \frac{p - 1}{n + 2} \cdot \frac{p - 2}{n + 3} \cdots \frac{2}{n + p - 1} \cdot \frac{1}{n + p} J_{n+p,0}.$$

Or, on a

$$J_{n+p,0} = \int_0^1 z^{n+p} dz = \frac{1}{n + p + 1}$$

de sorte qu'il vient finalement

$$J_{n,p} = \frac{p (p - 1) \ldots 2 \cdot 1}{(n + 1)(n + 2) \ldots (n + p + 1)} = \frac{n!\, p!}{(n + p + 1)!}.$$

La probabilité élémentaire est donc

$$P = z^n (1 - z)^p \frac{(n + p + 1)!}{n!\, p!} dz. \tag{1}$$

Nous posons, y étant très petit :

$$z = t - y.$$

On obtient aisément une valeur approchée de $\log P$ par un calcul analogue à celui qui a été développé au n° **28**. Nous omettrons ce calcul et nous contenterons d'en indiquer le résultat : en posant :

$$y(n - p)\sqrt{\frac{n}{2np}} = \lambda \qquad (3)$$

on obtient, comme première approximation, la valeur suivante de P,

$$P = \frac{1}{\sqrt{\pi}} e^{-\lambda^2} d\lambda \qquad (4)$$

c'est-à-dire que l'on retrouve la formule de Gauss, l'unité d'écart étant définie par la relation (3). La probabilité demandée dans l'énoncé est donc égale à $\Theta(\lambda)$, λ étant lié à y par la relation (3).

Quelque petit que soit y, la valeur de λ définie par cette relation (3) augmente lorsque les nombres n et p augmentent en conservant le même rapport : la probabilité pour que x soit compris entre $t - y$ et $t + y$ se rapproche donc indéfiniment de l'unité. Ainsi, lorsque les épreuves deviennent très nombreuses, on peut, comme il est naturel, considérer leur résultat global comme fournissant une indication assez exacte de la probabilité inconnue.

Donnons un exemple numérique, en rappelant que les calculs précédents ont été faits en supposant l'égalité des probabilités *à priori*.

Supposons que l'on ait :

$$n = 800$$

$$p = 100$$

La formule (3) donne alors

$$\lambda = 67,5\, y$$

et, en supposant $y = 0{,}05$, la valeur de $\theta\lambda$ diffère de l'unité d'un peu plus d'un dix-millième, car on a :

$$\theta(2{,}7) = 0{,}99986\ldots$$

On a d'ailleurs ici

$$\frac{n}{n+p} = \frac{8}{9} = 0{,}888\ldots$$

et l'on peut par suite conclure que la valeur de la probabilité d'amener face est comprise entre $0{,}859$ et $0{,}929$: cette conclusion a du moins une probabilité égale à $0{,}99986\ldots$

Si l'on avait supposé :

$$n = 80\,000$$
$$p = 10\,000$$

on aurait obtenu

$$\lambda = 6{,}5\,y$$

et par suite, en prenant $y = 0{,}005$, c'est-à-dire dix fois plus petit, on aurait été conduit à la même probabilité $\theta\lambda$; on aurait donc conclu que la probabilité d'amener face est comprise entre $0{,}885$ et $0{,}893$. On voit que la multiplication des expériences resserre les limites : la différence entre ces limites varie, comme il résulte du calcul précédent, en raison inverse de la racine carrée du nombre des expériences.

Si l'on convient de dire que la *précision* est inversement proportionnelle à la différence entre les limites obtenues, on pourra énoncer le résultat suivant : *la précision est proportionnelle à la racine carrée du nombre des expériences* (voir n° 55).

68. Vraie valeur d'une quantité mesurée. — Lorsque l'on a noté les résultats de diverses mesures d'une même quantité, le problème de la détermination de la valeur de cette quantité est un problème de probabilités des causes, car cette valeur est évidemment la cause principale qui a déterminé les valeurs obtenues pour les mesures. Si nous admettons les résultats du chapitre IX, nous pourrons, étant donnée une mesure, évaluer

la probabilité pour que cette mesure ait été obtenue, en supposant que la quantité à mesurer ait une valeur donnée quelconque. On peut ainsi arriver à trouver la probabilité a posteriori, c'est-à-dire en fonction des mesures, pour que la quantité à mesurer ait réellement une valeur donnée, en supposant connue la probabilité a priori. Dans le cas où l'on suppose uniformes les probabilités a priori, on retrouve ainsi par un calcul facile, la règle de la moyenne arithmétique, ainsi que les valeurs des probabilités pour que la vraie valeur présente un écart donné de la valeur fournie par la règle de la moyenne : ces probabilités suivent, comme on devait s'y attendre, la loi de Gauss.

Mais je n'insisterai pas sur ce point de vue, me contentant d'avoir signalé son intérêt théorique ; au point de vue des applications, il n'ajoute rien de nouveau aux résultats du chapitre IX.

CHAPITRE XIII

DÉTERMINATION DES CAUSES

69. Cas où la cause est inconnue. — Une forme fréquente sous laquelle se posent les problèmes de probabilités des causes est la suivante : *Tel résultat est-il dû au hasard, ou a-t-il une cause ?* On a souvent fait observer combien cet énoncé manque de précision : Bertrand a beaucoup insisté sur ce point : nous aurons l'occasion de citer les plus intéressantes de ses observations. Mais tout ce que l'on peut objecter au point de vue logique ne saurait empêcher la question précédente de se poser d'elle-même dans bien des cas : la théorie des probabilités ne peut donc refuser de l'examiner et d'y faire une réponse ; la précision de cette réponse sera naturellement limitée par le manque de précision de la question ; mais refuser de répondre sous prétexte que la réponse ne peut pas être absolument précise, c'est se placer sur un terrain purement abstrait et méconnaître le caractère qu'ont forcément les applications des mathématiques. Le calcul fournit, il est vrai, une réponse précise à toute question précise ; mais, pratiquement, il n'y a jamais de question précise : les données expérimentales comportent forcément un certain jeu ; la même imprécision affecte donc le résultat du calcul et la soi-disant précision théoriquement absolue de ce calcul n'est qu'une pure illusion.

Comme exemples de questions dans lesquelles la cause est inconnue, ou dans lesquelles *le problème est précisément de savoir s'il y a une cause*, on peut citer un grand nombre de

recherches de psychologie expérimentale [1] : par exemple, on constate chez un certain nombre d'individus, la coïncidence entre une particularité physique et une particularité de nature psychologique ; dans quelles limites peut-on en conclure à une corrélation réelle ?

Signalons aussi les questions fort nombreuses que l'on peut se poser sur la langue d'un auteur : chaque écrivain a pour la longueur des mots et des phrases, des habitudes particulières que l'on peut traduire en nombres ; de telles remarques peuvent mettre sur la voie d'interpolations de textes. Dans d'autres cas, on pourra se poser et résoudre par la méthode statistique des questions de métrique [2].

On a aussi appliqué les méthodes de la théorie des probabilités à des questions d'art : une école de sculpture, par exemple, est caractérisée par la fixité de certains rapports entre les mesures de diverses parties du corps ; des écarts numériques peuvent faire présumer une différence d'école que confirme une étude proprement esthétique [3].

Dans toutes ces questions, le calcul est surtout destiné à fournir des présomptions qu'il est nécessaire de mettre à l'épreuve d'une critique directe du problème posé, par la technique propre de son art.

Nous laisserons de côté ces questions et bien d'autres, malgré leur intérêt, pour nous borner à donner quelques détails sur les applications à des sujets plus proprement scientifiques, c'est-à-dire se rapportant à des sciences considérées comme telles depuis plus longtemps.

70. La distribution des étoiles sur la sphère céleste. — Cette question est une des plus passionnantes de la philosophie

---

[1] Voir mon étude sur *Le calcul des probabilités et la méthode des majorités* dans l'Année Psychologique, t. XIV, 1908 pp. 125-151.

[2] Les difficultés de ce genre de questions, et en même temps l'utilité réelle de la théorie des probabilités pour les résoudre ont été signalées avec beaucoup de justesse par M. Daniel Serruys dans une note trop technique pour que je puisse l'analyser ici : Les procédés toniques d'Himérius et les origines du « Cursus » byzantin (Mélanges Louis Havet : Philologie et linguistique. Hachette, 1909).

[3] Voir Jean Laran : La méthode statistique dans un problème d'archéologie (*Revue du Mois* du 10 avril 1907, t. III, p. 433).

naturelle, car elle touche aux origines et aux destinées de notre univers. Sauf quelques exceptions relativement rares, nous ignorons les distances des étoiles : une première difficulté se présente donc : lorsque deux ou plusieurs étoiles sont rapprochées sur la carte du ciel, peut-on en conclure qu'elles sont rapprochées dans l'espace ? Il ne saurait être question d'apporter à cette question une réponse rigoureuse ; mais l'étude des probabilités peut, dans certains cas, donner de très fortes présomptions en faveur de l'affirmative. A moins d'admettre en effet, des alignements réguliers analogues à ceux de la théorie des réseaux cristallins, alignements dont aucune raison n'autorise à soupçonner l'existence, il est clair que, si deux ou plusieurs étoiles sont extrêmement éloignées dans l'espace, les probabilités pour que leurs positions sur la sphère céleste soient rapprochées les unes des autres se calculeront d'après les principes du n° **45**. Si dès lors, on fait entrer en ligne de compte le nombre des étoiles de grandeur déterminée, par exemple le nombre des étoiles visibles à l'œil nu, ou avec telle lunette, on pourra calculer quelle est la probabilité pour qu'un nombre déterminé de ces étoiles soient situées à l'intérieur d'un petit cercle donné de la sphère céleste. S'il existe réellement un tel groupement dont la probabilité ainsi calculée soit très faible, on sera autorisé à penser que ce groupement a une *cause* autre que le hasard, c'est-à-dire qu'il est formé d'étoiles réellement rapprochées dans l'espace.

Bertrand fait à ce genre de raisonnements des objections qui ne me paraissent pas convaincantes :

« Les Pléiades, dit-il [1], semblent plus rapprochées les unes « des autres qu'il n'est *naturel*. L'assertion est digne d'intérêt ; « mais, si l'on veut traduire la conséquence en chiffres, les élé- « ments font défaut. Faut-il, pour préciser cette idée vague de « rapprochement, chercher le plus petit cercle qui contienne le « groupe ? la plus grande des distances angulaires ? la somme des « carrés de toutes les distances ? l'aire du polygone sphérique « dont quelques-unes des étoiles sont les sommets et qui con- « tient les autres dans son intérieur ? Toutes ces grandeurs,

[1] Bertrand. — *Calcul des Probabilités*, p. 170.

« dans le groupe des Pléïades, sont plus petites qu'il n'est vrai-
« semblable. Laquelle d'entre elles donnera la mesure de
« l'invraisemblance? Si trois étoiles forment un triangle équila-
« téral, faut-il faire entrer cette circonstance, assurément peu
« probable a priori, au nombre de celles qui révèlent une
« cause? »

Ce que l'on traduit en chiffres, c'est la probabilité pour qu'un groupement soit dû au hasard : quoiqu'en dise Bertrand, le résultat numérique ainsi défini ne dépendra que dans une très faible mesure de la forme particulière donnée à la définition du groupement : plus petit cercle le renfermant, ou polygone sphérique convexe, etc. Pratiquement, on choisira celle de ces définitions qui conduira aux calculs les plus simples. Ces calculs conduiront à un résultat tel que le suivant : le hasard produirait une fois sur 23 000 un tel groupement. Une fois ce résultat énoncé, la théorie des probabilités a donné tout ce qu'elle pouvait donner et doit céder la place à l'induction et à l'hypothèse, que l'on cherchera ultérieurement à vérifier par d'autres méthodes, si c'est possible.

Disons un mot de la réflexion de Bertrand relativement au triangle équilatéral que formeraient trois étoiles : elle se rattache à la question du nombre rond. Si l'on considère un nombre pris au hasard entre 1 000 000 et 2 000 000 la probabilité pour qu'il soit égal à 1 342 517 est égale à un millionième ; la probabilité pour qu'il soit égal à 1 500 000 est aussi égale à un millionième. On considérera cependant volontiers cette dernière éventualité comme moins probable que la première ; cela tient à ce qu'on ne se représente jamais *individuellement* un nombre tel que 1 542 317 : on le regarde comme le *type* de nombres d'apparences analogues et si, en le transcrivant, on modifie un chiffre, on s'en aperçoit à peine et l'on ne distingue pas 1 324 519 de 1 324 517 : le lecteur a besoin de faire un effort pour s'assurer que les quatre nombres écrits dans les lignes précédentes sont tous différents.

Lorsque l'on a observé un nombre tel que le précédent comme évaluation d'un angle en dixièmes de secondes centésimales, on ne songe pas à se poser la question de savoir quelle était la probabilité pour que cet angle fût précisément égal à 13° 42' 51'' 7 ;

car on ne se serait jamais posé cette question précise avant d'avoir mesuré l'angle. Il faut bien que cet angle ait une valeur et, quelle que soit sa valeur à un dixième de seconde près, on pourrait, après l'avoir mesurée, dire que la probabilité *a priori*, pour que cette valeur soit précisément telle qu'elle est, est un dix-millionième, et que c'est là un fait bien extraordinaire. C'est un sophisme de ce genre qui viciait un système d'expertise fondé sur la mensuration des écritures, système qui eut un moment de célébrité il y a quelques années, pour des raisons complètement étrangères aux mathématiques.

La question est de savoir si l'on doit faire ces mêmes réserves dans le cas où l'on constate qu'un des angles du triangle formé par trois étoiles a une valeur *remarquable* et est, par exemple, égal à l'angle du triangle équilatéral (66°66'66",6), ou à un demi-angle droit : 50 grades, ou 5 000 000 dixièmes de seconde. Voici ce que l'on peut dire à ce sujet : on doit se défier beaucoup de la tendance que l'on a à regarder comme *remarquable* une circonstance que l'on n'avait pas précisée *avant l'expérience*, car le nombre des circonstances qui peuvent apparaître comme remarquables, à divers points de vue, est très considérable [1].

Or la valeur de l'angle déterminé par les positions apparentes de trois étoiles est une fonction compliquée des positions réelles des étoiles et de la position de notre système solaire : le fait que cet angle serait précisément égal à 6 666 666 dixièmes de seconde centésimale ne doit pas paraître plus remarquable que le fait qu'il serait égal précisément à 1 234 567 dixièmes de seconde, ce qui est aussi un nombre « remarquable », ou à

(1) Voici un exemple simple de ce fait ; considérons un nombre de quatre chiffres tel que 2545 et supposons qu'on le décompose par la pensée en deux tranches de deux chiffres 25 et 45 ; on remarque ici que ces deux nombres se terminent par le même chiffre 5, ce qui est une *circonstance particulière* ; de même pour 2525, les deux nombres 25 et 25 commencent par le même chiffre 2 ; pour 2550 les nombres 25 et 50 sont symétriques ; pour 2250, le second nombre 50 est double de 25, pour 2536, 25 et 36 sont les carrés de deux entiers consécutifs, etc. On en arrive ainsi aisément à trouver à près de la moitié des nombres de quatre chiffres *quelque chose de particulier*, c'est là un petit jeu auquel on peut s'amuser à Paris en regardant passer les fiacres, qui ont en général un numéro de 4 chiffres ; on serait porté à croire, si l'on ne précisait pas, que les numéros *particuliers* sont plus fréquents que ne le voudrait la théorie des probabilités et à chercher la cause de cette fréquence.

2 385 256, ce qui est un nombre ne présentant à priori rien de particulier.

Au contraire, la question du groupement dans l'espace est une de celles qu'il est naturel de se poser a priori : la théorie des probabilités ne permet pas de la résoudre avec certitude, mais précise les conséquences que l'on peut tirer des observations.

Cette précision n'est pas inutile, car elle remplace souvent d'une manière heureuse les inductions trop rapides du « bon sens ». Sans doute, il n'est besoin d'aucun calcul pour savoir que la voie lactée est un amas exceptionnel d'étoiles, et le calcul n'ajouterait pas grand'chose à cette impression. D'autre part, il est des cas où l'on sera conduit à ce résultat que la probabilité d'un groupement est un tiers ou un quart : la présomption pour qu'il y ait une cause est ici trop faible pour que le résultat ait un intérêt : il ne prouve rien, pas plus que le fait qu'un joueur d'écarté a retourné une fois le roi ne prouve qu'il est un tricheur. Mais, entre ces deux cas extrêmes, celui où la probabilité du groupement par le hasard est tellement faible qu'on a l'intuition qu'elle est pratiquement nulle et celui où cette probabilité est assez élevée pour qu'on ne puisse rien conclure, il est bien d'autres cas, ceux où la probabilité est comprise entre un centième et un cent millième, par exemple : le calcul de la valeur qu'elle a entre ces limites n'est pas inutile et le renseignement qu'il fournit peut être précieux, quoi qu'en ait pu penser Bertrand.

Je me suis étendu un peu longuement sur ces réflexions, afin de n'avoir pas à y revenir dans les autres problèmes que nous allons maintenant passer en revue.

71. Les valeurs des poids atomiques. — On sait que les chimistes ont remarqué depuis longtemps que les valeurs des poids atomiques sont le plus souvent des nombres entiers. Ce fait longtemps admis a été sérieusement contesté, à la suite de mesures plus précises ; actuellement, les valeurs admises par la Commission internationale des poids atomiques ne sont généralement pas entières.

Il subsiste néanmoins entre elles des relations remarquables ; pour ne citer qu'un exemple parmi les corps les mieux connus,

le rapport des poids atomiques du carbone et de l'oxygène est, à l'approximation des erreurs de mesure, égal au rapport de 3 à 8.

Il est dès lors naturel de rechercher *si les rapports des poids atomiques entre eux ne sont pas plus voisins de nombres commensurables simples que s'ils étaient choisis au hasard*. Il ne semble pas que la question posée sous cette forme précise ait été étudiée avec le soin que mériterait son intérêt.

On peut rattacher cette question à la théorie des fractions continues et des probabilités dénombrables : on est conduit, comme je l'ai indiqué dans mon cours de 1908-1909, à développer en fraction continue les rapports des poids atomiques considérés deux à deux, et à rechercher si les quotients incomplets prennent les diverses valeurs entières plus ou moins souvent que ne l'exigerait le hasard. Comme les combinaisons 2 à 2 des corps simples sont au nombre de plusieurs milliers, on possède un assez grand nombre de résultats pour utiliser avec fruit la théorie des probabilités. La principale difficulté de cette recherche est la limitation de la précision avec laquelle chaque rapport est connu ; cette limitation exigerait évidemment la discussion complète de toutes les expériences par lesquelles les deux poids atomiques dont on étudie le rapport sont reliés l'un à l'autre : c'est là un travail considérable, mais qui me paraîtrait justifié par l'intérêt qui s'attache au problème ainsi posé.

Son étude conduirait, vraisemblablement autant que j'en puis juger d'après les calculs que j'ai faits à la conclusion que les poids atomiques, de certains groupes de corps simples tout au moins, diffèrent très peu de nombres ayant entre eux des rapports simples ; ces nombres pourraient être appelés *poids atomiques virtuels* et il resterait à rechercher quelle est leur signification physique et quelles sont les raisons pour lesquelles les poids atomiques expérimentaux s'en distinguent par de petites différences.

On sait que c'est un fait fréquent dans les sciences physiques que l'existence d'une petite différence résiduelle entre les valeurs mesurées expérimentalement et certaines valeurs virtuelles que l'on peut définir théoriquement et qui suivent des lois simples.

72. Les applications biométriques. — On a donné récemment le nom de *Biométrique* à l'ensemble des recherches dans lesquelles sont utilisées des ensembles de mesures effectuées sur des êtres vivants. Je n'ai pas à faire ici l'histoire de la biométrique, pour laquelle je renverrai à une substantielle étude de M. Vito Volterra [1] ; qu'il me suffise de citer le nom de Quételet, qui fut un précurseur et ceux des anglais Francis Galton et Karl Pearson.

Les problèmes que se pose la biométrique sont de natures diverses ; indiquons deux des plus importants : l'homogénéité des groupes et les corrélations.

Supposons que l'on mesure la taille d'un grand nombre de Français adultes ; si l'on convient de considérer la valeur moyenne des mesures comme la valeur exacte que devrait avoir la taille d'un Français, on constate que les « erreurs », c'est-à-dire les différences positives ou négatives entre cette valeur théorique et la valeur réelle se répartissent précisément suivant la loi de Gauss ; tout se passe comme si les Français avaient tous la même taille, égale à la moyenne, mais étaient mesurés par un expérimentateur très maladroit, dont les erreurs de mesure suivraient la loi de Gauss. Pour employer le langage défini au n° 53, nous pourrons dire que l'ensemble des tailles des Français constitue une série *normale*. Biologiquement, ceci coïncide avec le fait que les Français constituent un groupe biologique suffisamment homogène.

Supposons que nous fassions la même expérience dans une ville dont les habitants appartiennent à deux races différentes, les uns étant blancs et les autres jaunes. Si l'on mesure tous les habitants adultes et que l'on note les mesures sans distinguer les races, on constate que la moyenne n'a plus la même signification et que le groupement autour de la moyenne ne suit pas de loi simple telle que celle de Gauss. Il y a superposition de deux phénomènes distincts, dont chacun suit la loi de Gauss, mais dont la somme ne suit pas cette loi : les tailles des blancs

---

[1] Vito Volterra. Sur l'application des mathématiques aux sciences biologiques et sociales (*Revue du Mois* du 10 janvier 1906, t. I, p. 1).

se groupent autour de leur propre moyenne suivant la loi de Gauss et de même les tailles des jaunes autour de leur moyenne : la moyenne générale n'a aucune signification : en particulier, il peut très bien arriver que cette moyenne générale ne soit pas la taille la plus fréquente. Ceci correspond au fait biologique que les races sont distinctes, et ce fait, s'il n'avait pas été connu, aurait pu être décelé par l'étude des moyennes.

Sans entrer dans des détails dont l'étude approfondie mériterait à elle seule tout un ouvrage, on conçoit comment l'étude expérimentale de faits analogues a pu conduire à la conclusion générale suivante : *le caractère biométrique intrinsèque des séries normales est la pureté de la race*. Je laisse de côté les commentaires nombreux dont aurait besoin un énoncé aussi général : bien des idées auraient à être précisées, ce qui est l'affaire du biologiste encore plus que du mathématicien : néanmoins, telle qu'elle est, cette théorie a prouvé sa valeur scientifique en se rendant utile par des applications pratiques. Je ne puis les mentionner toutes : je citerai toutefois les travaux très intéressants sur les orges de brasserie qui ont été inaugurés par M. Hjalmar Nilsson au laboratoire de Svalöf en Suède, et acclimatés en France par M. Blaringhem, dont les recherches personnelles ont été des plus intéressantes. On arrive, par l'application du caractère des races pures, à sélectionner et à cultiver dans des régions étendues des races d'orge extrêmement pures, dont la fixité est très appréciée par les brasseurs qui les emploient : on sait quels sont les avantages industriels de l'homogénéité de la matière première.

On voit combien est grande l'importance de la théorie biométrique de l'homogénéité ; la théorie des corrélations n'est pas moins intéressante, car elle peut donner la clef des principaux problèmes de l'hérédité. On effectue diverses mesures sur plusieurs individus d'une même race, sur leurs parents et sur leurs collatéraux ; quelles sont les corrélations de ces diverses mesures entre elles ? Par exemple, connaissant la taille d'un individu, peut-on en conclure quelque chose relativement à la dimension de l'avant-bras de son frère ? On voit que le concept de *cause* s'élargit ici singulièrement ; on ne dira habituellement pas que la taille élevée d'un individu est la *cause* de la longueur de l'avant-

bras de son frère : cette extension de la notion de cause est parfaitement légitime au point de vue de la théorie des probabilités [1]. Je me bornerai, pour l'étude de cette théorie des corrélations, à renvoyer aux travaux de Karl Pearson et de ses élèves, parus notamment dans les *Proceedings of the Royal Society of London* et dans le journal spécial *Biometrika*.

73. Conclusion. — On voit combien est vaste le champ des applications de la théorie des probabilités : encore ne les avons-nous pas mentionnées toutes et avons-nous dû passer trop rapidement sur beaucoup d'entre elles. La théorie des probabilités est la science des collectivités et, dans bien des questions, biologiques ou sociales notamment, la tendance actuelle est de substituer l'étude des collectivités à celle des individus. Cette substitution est souvent délicate, car il s'agit de savoir si la considération de la collectivité comme un tout sensiblement homogène est légitime : la théorie des probabilités donne souvent des indications précieuses sur cette question préjudicielle, qui est fondamentale.

Je tiens à rappeler en terminant une remarque que j'ai déjà faite plusieurs fois, mais sur laquelle on ne saurait trop insister : la théorie des probabilités, comme toute théorie mathématique d'ailleurs, ne saurait prétendre résoudre *à priori* des questions concrètes, qui sont du domaine de l'expérience : son seul rôle — déjà très beau si elle sait le remplir — est de guider l'expérience et l'observation par l'interprétation qu'elle fournit de leurs résultats.

---

[1] Dans une conception entièrement déterministe de l'univers, il n'y a pas de *cause* au sens habituel que l'on donne à ce terme : l'ensemble de l'univers doit, à tout instant, être regardé comme la cause de tous les évènements passés, présents ou futurs ; il n'est pas en effet possible de modifier un seul phénomène sans modifier tous les autres, car il faut pour cela concevoir *un autre* univers, celui dans lequel nous vivons ne pouvant, dans cette hypothèse, être conçu autre qu'il n'est. On doit donc, à ce point de vue entièrement déterministe, remplacer la notion de cause par celle de corrélation, qui s'applique également quel que soit l'ordre de succession des phénomènes étudiés et leurs relations directes apparentes.

TABLE DES MATIÈRES

PRÉFACE v

LIVRE PREMIER

PROBABILITÉS DISCONTINUES

CHAPITRE PREMIER

LE JEU DE PILE OU FACE

Pages

1. Définition du jeu 3
2. Étude des cas simples 4
3. Remarques sur les cas simples 7
4. Triangle arithmétique de Pascal. 8
5. Quelques problèmes élémentaires 10
6. Espérance mathématique 12
7. Problèmes sur l'espérance mathématique 14
8. Formules générales pour n parties 16
9. Remarques sur quelques paradoxes. 18

CHAPITRE II

QUELQUES DÉFINITIONS ET QUELQUES THÉORÈMES

10. Définition de la probabilité 21
11. Propriétés essentielles de la probabilité 22
12. Probabilités totales 25
13. Probabilités composées 27
14. Application aux tirages successifs d'une même urne 30
15. Autres problèmes. 33

CHAPITRE III

FORMULES D'APPROXIMATION

16. But de ce chapitre 38
17. Formule de Stirling. Première approximation 39
18. Application à l'égalité des joueurs à pile ou face 39
19. Deuxième approximation 42
20. Résumé des formules 44

CHAPITRE IV

ETUDE APPROFONDIE DU JEU DE PILE OU FACE

21. Formule fondamentale 46
22. Unité d'écart 49
23. Emploi de la variable continue 51
24. Définition de la fonction $\Theta(\lambda)$. 54
25. Problèmes où on utilise $\Theta(\lambda)$. 55
26. Ecart le plus probable, écart probable, écart médian 56
27. Valeur moyenne d'une fonction quelconque 58

CHAPITRE V

LOI DES GRANDS NOMBRES ET LOI DES ÉCARTS

28. Etude des épreuves répétées dans le cas général 60
29. Loi des grands nombres 63
30. Théorème de Bernoulli 64
31. Quelques problèmes 65
32. Cas où la probabilité a deux valeurs distinctes 70
33. Substitution d'un seul groupe d'épreuves à deux groupes . . 73
34. Extension au cas de plusieurs groupes 76
35. Résumé des résultats obtenus 78

LIVRE II

PROBABILITÉS CONTINUES

CHAPITRE VI

DÉFINITION DE LA PROBABILITÉ GÉOMÉTRIQUE

36. Position d'un point sur un segment de droite 83
37. Objection à la définition 84
38. Position d'un point dans le plan et dans l'espace 86
39. Autres probabilités géométriques 87

CHAPITRE VII

QUELQUES PROBLÈMES DE PROBABILITÉS GÉOMÉTRIQUES

40. Remarque préliminaire. 90
41. Une difficulté de la vérification expérimentale 91
42. Problèmes relatifs à la position de points sur une droite . . 93
43. Problèmes relatifs à la position de points sur un cercle . . . 98
44. Problème relatif à la position de points dans un aire plane. . 99
45. Problèmes relatifs à la position de points sur la sphère . . . 100
46. Autres problèmes. 104

CHAPITRE VIII

INTRODUCTION DES FONCTIONS ARBITRAIRES

47. Equations générales. 114
48. Cas où le résultat est indépendant du choix de la fonction arbitraire 117
49. Application à la distribution des petites planètes 122
50. Application à la théorie cinétique des gaz 124

CHAPITRE IX

ERREURS D'OBSERVATION — LOI DE GAUSS

71. Détermination du problème 129
72. Énoncé de la loi de Gauss 130
73. Développement de la loi de Gauss 132
74. Inconvénient de la loi de Gauss 134
75. Précision d'une série de mesures 135
76. Poids d'un résultat. Moyennes 137

LIVRE III

PROBABILITÉS DES CAUSES

CHAPITRE X

CAS DES PROBABILITÉS DISCONTINUES

77. Quelques problèmes simples 143
78. Formule de Bayes 145
79. Applications diverses 153

CHAPITRE XI

PROBLÈMES STATISTIQUES

80. Notion de la probabilité statistique 158
81. Naissances masculines et féminines 158
82. Étude des accouchements doubles 160
83. Naissances dans une même famille 163
84. Décès 165
85. Sélection des unes 167

CHAPITRE XII

CAS DES PROBABILITÉS CONTINUES

86. Étude d'un problème simple 170
87. Cas particuliers 171
88. Vraie valeur d'une quantité mesurée 175

CHAPITRE XIII

DÉTERMINATION DES CAUSES

69. Cas où la cause est inconnue 177
70. La distribution des étoiles sur la sphère céleste 178
71. Les valeurs des poids atomiques. 182
72. Les applications biométriques 184
73. Conclusion 186

SAINT-AMAND, CHER. — IMPRIMERIE BUSSIÈRE

QA
273
B66

Borel, Émile Felix Édouard Justin
Éléments de la théorie des probabilités

Physical &
Applied Sci.

PLEASE DO NOT REMOVE
CARDS OR SLIPS FROM THIS POCKET

UNIVERSITY OF TORONTO LIBRARY